Libros de Cátedra

Algebra Lineal con Aplicaciones

Parte I

Raúl Rossignoli (coordinador)

FACULTAD DE
INGENIERÍA

exactas





Algebra Lineal con Aplicaciones

Parte I

V. Costa, R. Rossignoli,
C. Sorichetti y V. Vampa

Coordinador: R. Rossignoli

Facultad de Ingeniería





Dedicatoria

En recuerdo de la Profesora Nélida Echebest.

Prólogo

Este libro está pensado como texto para ser utilizado en la parte inicial de un curso, de duración semestral, sobre Algebra Lineal para carreras de Ingeniería y otras Ciencias aplicadas. El libro está basado en las guías teórico-prácticas elaboradas inicialmente por la que fuera Profesora Titular de la asignatura Matemática C de la Facultad de Ingeniería de la UNLP, Lic. Nélida Echebest. Esta base fue luego reelaborada y enriquecida con aportes de los presentes autores, profesores de dicha asignatura, teniendo como referencia la bibliografía [1–8] indicada al final del presente libro. Dicha asignatura, correspondiente al tercer trimestre de las carreras de Ingeniería de la Universidad Nacional de La Plata, introduce herramientas básicas que son de utilidad en la modelización y resolución de problemas de Ingeniería, Física, Química, etc. Por esta misma razón, el presente libro puede resultar también útil para cursos destinados a estudiantes de otras disciplinas científicas. Se ha dado por supuesto que el lector ha adquirido, previamente, una formación básica sobre Análisis Matemático en una y varias variables reales. El libro contiene desarrollos teóricos, incluyendo las principales demostraciones, y además numerosos ejemplos resueltos en detalle, junto con interpretaciones geométricas y figuras, para reforzar y clarificar los conceptos introducidos. Asimismo, se presenta una amplia variedad de problemas y aplicaciones.

El capítulo I, Sistemas de Ecuaciones Lineales, introduce las técnicas básicas para resolver estos sistemas con un número arbitrario de ecuaciones e incógnitas. Se describe en detalle el método de eliminación de Gauss y se determinan las condiciones para las que el sistema resulta compatible determinado (solución única), compatible indeterminado e incompatible.

El capítulo II, Matrices, introduce las operaciones matriciales básicas, para luego focalizarse en la representación matricial de sistemas lineales. Se introduce también el concepto de matriz inversa y matriz singular, y se incluyen algunas aplicaciones.

El capítulo III, Determinantes, introduce gradualmente el concepto de determinante, vinculándolo con la resolución de sistemas de n ecuaciones con n incógnitas y las condiciones que aseguran solución única. También se pone énfasis en su interpretación geométrica, sus propiedades fundamentales y su evaluación eficiente.

En el capítulo IV, se define el concepto de Espacio Vectorial, extendiendo a espacios vectoriales generales abstractos las nociones básicas de suma de vectores y multiplicación por un escalar en el plano y el espacio tridimensional, que supondremos ya conocidas por el lector. Se presentan en forma detallada los conceptos de subespacio, independencia lineal, base y dimensión, incluyendo la noción general de coordenada y cambio de base. Luego se aplican estos conceptos para retomar, desde una perspectiva más amplia, los sistemas de ecuaciones lineales generales y la caracterización del conjunto solución, estudiados previamente, relacionándolos con las propiedades de la matriz correspondiente y de los espacios vectoriales asociados a sus filas y columnas.

Finalmente, en el capítulo V se define el concepto de Transformación Lineal entre espacios vectoriales generales, haciendo primero hincapié en aquellas transformaciones relacionadas con operaciones geométricas simples en el plano y el espacio. Se discuten sus

propiedades fundamentales, y se pone especial énfasis en la representación matricial de las mismas. También se retoman los sistemas lineales desde esta perspectiva, abarcando así todos los conceptos discutidos en los capítulos previos.

De esta forma, esta primera parte proporciona la base para los conceptos que se estudiarán en la Parte II, entre los que se incluyen autovalores y diagonalización, ortogonalización, ecuaciones diferenciales lineales y sistemas de ecuaciones diferenciales lineales, además de utilización de software.

Agradecimientos

A nuestros colegas y alumnos de la Facultad de Ingeniería, que enriquecieron el contenido de esta obra con sus enseñanzas, preguntas y sugerencias. A nuestro colega Prof. Alejandro Mesón, por su aporte en el capítulo de matrices.

A la Universidad Nacional de La Plata y la Facultad de Ingeniería, por haber apoyado la propuesta de realizar este libro y brindar la oportunidad de publicarlo. Este trabajo se enmarca también en el Proyecto Acreditado en el Programa de Incentivos de la UNLP, *Diseño, implementación y análisis de estrategias didácticas en Ciencias Básicas en carreras de Ingeniería*, en el que participan V. Costa y R. Rossignoli. R. Rossignoli agradece también a la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires.

Índice general

Capítulo 1

Sistemas de Ecuaciones Lineales

1.1. Introducción

El objetivo básico de este capítulo es mostrar una metodología general y eficiente para resolver sistemas de ecuaciones *lineales*, válida para cualquier número de ecuaciones y de incógnitas. La metodología permitirá, en primer lugar, determinar si el sistema es *compatible*, es decir, si tiene solución. Luego proporcionará una manera eficiente de obtener *todas* las soluciones posibles. Veremos entonces que los sistemas lineales compatibles pueden ser de dos tipos: *determinados*, que son aquellos que poseen *solución única*, e *indeterminados*, que son aquellos que poseen *infinitas soluciones* (asumiendo que las incógnitas pueden tomar cualquier valor real). Estos últimos tendrán un conjunto de variables *libres* (o independientes), que determinarán el conjunto de soluciones.

Los sistemas de ecuaciones lineales son aquellos que involucran sólo la potencia 1 de las variables incógnitas (y sólo sumas de estas variables multiplicadas por constantes). Son de uso común y frecuente en matemática, ingeniería y las ciencias en general, siendo los más fáciles de resolver (y más antiguos: sistemas de simples (2×2) de ecuaciones lineales eran resueltos ya en la antigua Babilonia). Veamos primero algunos ejemplos simples.

1) **Palanca en equilibrio**

Comencemos con un problema básico de Física: una palanca en equilibrio. Supongamos que se tienen tres objetos A, B y C, uno con peso conocido (por ejemplo C). Se desea conocer el peso de los otros dos objetos. Como dato, se sabe que se ha logrado el equilibrio en las dos configuraciones siguientes (distancias en metros):

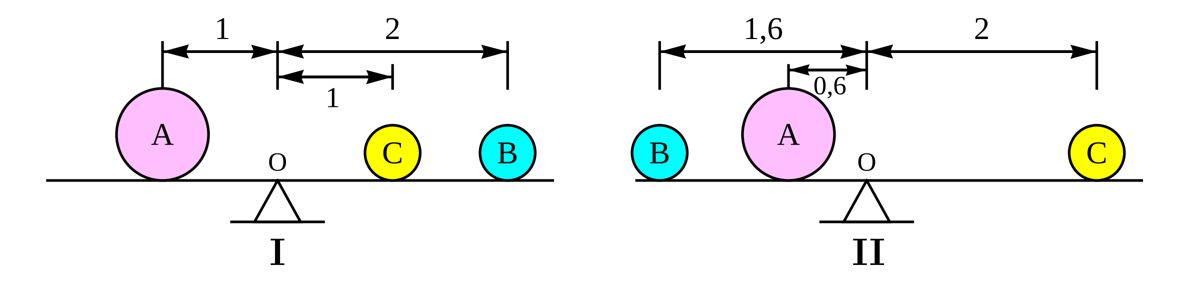


Figura 1.1: Palancas en equilibrio.

Considerando que en un sistema en equilibrio la suma de los momentos de las fuerzas aplicadas respecto a un punto cualquiera, por ejemplo el punto de apoyo O, debe ser 0, obtenemos las ecuaciones:

$$\left\{ \begin{array}{rcl} P_A - 2\,P_B - P_C & = & 0 \\ 0,6\,P_A + 1,6\,P_B - 2\,P_C & = & 0 \end{array} \right. \qquad (1.1)$$

donde P_A, P_B, P_C denotan los pesos de los objetos. Si P_C es conocido, este es un sistema de dos ecuaciones lineales con dos incógnitas: P_A y P_B. Por ejemplo, si $P_C = 2kg$ y expresamos P_A y P_B también en kg, obtenemos el sistema

$$\left\{ \begin{array}{rcl} P_A - 2P_B & = & 2 \\ 0,6P_A + 1,6P_B & = & 4 \end{array} \right. \qquad (1.2)$$

Este sistema tiene solución única $P_A = 4kg$, $P_B = 1kg$, como se puede comprobar fácilmente. Por el contrario, si P_C es desconocido, el sistema (1.1) tendrá infinitas soluciones para P_A, P_B, P_C (con un sólo parámetro libre, como veremos), mientras que si P_C y P_B son ambos conocidos, siendo P_A la única incógnita, el sistema puede ser compatible o incompatible, dependiendo de los valores de P_B y P_C.

2) Flujo de redes

Una *red* consiste en un conjunto de puntos llamados *nodos*, con líneas o arcos que los conectan denominadas *ramas*. La dirección del flujo se indica en cada rama y la cantidad (o tasa) de flujo se denota por medio de una variable. El supuesto básico estándar en una red de flujos es que el flujo que entra a la red es el mismo que sale de la red, y que el flujo entrante en un nodo es igual al flujo saliente del nodo. Por ejemplo, en la figura siguiente se muestra una red elemental con un sólo nodo.

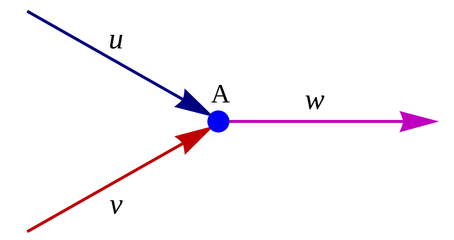


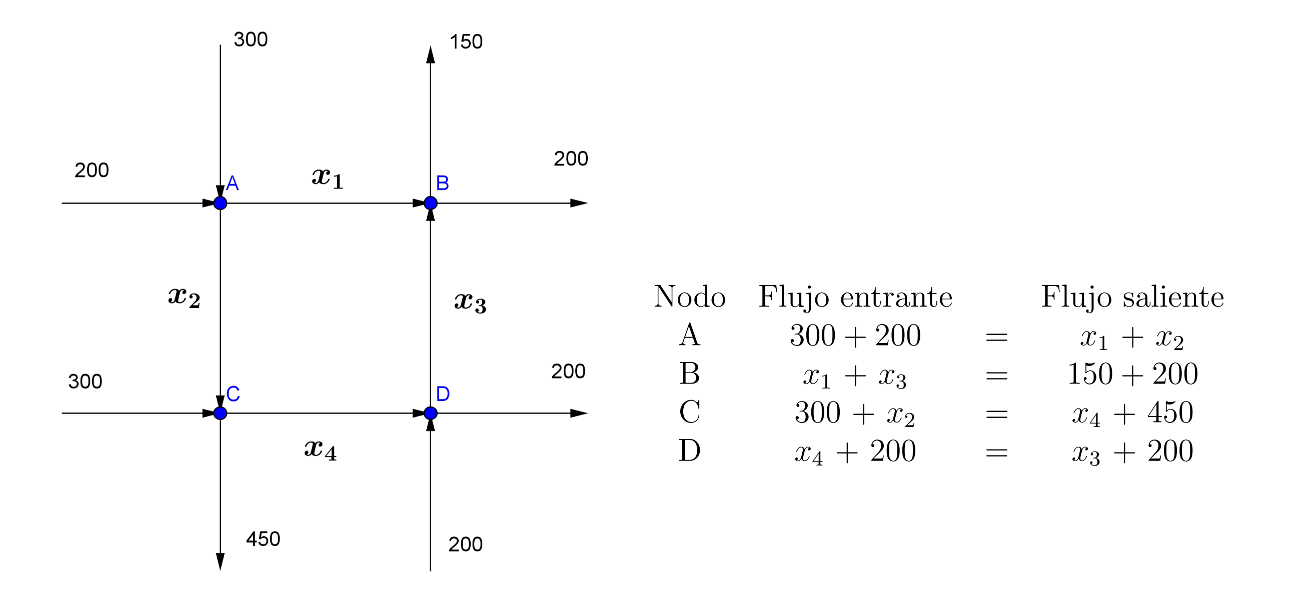
Figura 1.2: Esquema de red elemental con un nodo.

En este caso, los flujos entrantes u y v y el flujo saliente w deben satisfacer:

$$u + v = w \tag{1.3}$$

Múltiples problemas de ingeniería, ciencias sociales y naturales (entre otros) se pueden *modelar* a partir del planteo de un *flujo de redes*. Los flujos pueden ser de tráfico en una ciudad, de aviones en aeropuertos, de corriente en un circuito eléctrico, de distribución de mercaderías entre mayoristas y vendedores, de caudales en una red de tuberías, etc.

Por ejemplo, supongamos que en una cierta ciudad se va a realizar un arreglo en las calles y se quiere conocer el flujo de tránsito en alguna de ellas para tomar decisiones en cuanto a su redireccionamiento. En la red de la figura siguiente se indica el flujo de tráfico que entra o sale de cada calle, en número de vehículos por hora, considerando el tráfico promedio durante las horas pico. Se modela el problema. Identificamos los nodos: A, B, C y D, y los flujos a conocer: x_1, x_2, x_3 y x_4. Para cada nodo se debe verificar lo siguiente (flujo entrante, igual al flujo saliente):



Nodo	Flujo entrante		Flujo saliente
A	$300 + 200$	$=$	$x_1 + x_2$
B	$x_1 + x_3$	$=$	$150 + 200$
C	$300 + x_2$	$=$	$x_4 + 450$
D	$x_4 + 200$	$=$	$x_3 + 200$

Figura 1.3: Esquema de red.

Se obtiene así un sistema de 4 ecuaciones lineales con 4 incógnitas x_1, x_2, x_3, x_4. Como se comprobará luego, este sistema es compatible indeterminado. Pero si se conoce una de las variables x_i, resulta compatible determinado para las restantes.

3) Distribución de temperatura en estado estacionario en una placa plana

Un aspecto importante en el estudio de la *transferencia de calor* es determinar la distribución de temperatura en estado estable sobre una placa delgada cuando se conoce la temperatura en el borde. Supongamos que la placa mostrada en la figura representa la sección transversal de una viga de metal con un flujo de calor insignificante en la dirección perpendicular a la placa.

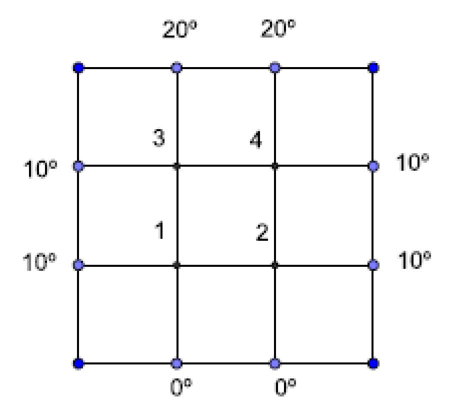


Figura 1.4: Modelización simple de la distribución de temperatura en un una placa plana.

El problema puede ser modelado de la siguiente forma. Sean T_1, T_2, T_3, T_4 las temperaturas en los cuatro nodos interiores en la malla que se muestra en la figura. En un nodo, la temperatura es aproximadamente igual al promedio de las temperaturas de los cuatro nodos más cercanos (a la izquierda, arriba, a la derecha y abajo). Es decir, para cada nodo, se obtienen las igualdades:

$$\begin{array}{rcl} T_1 & = & \frac{T_2+T_3+10+0}{4} \\ T_2 & = & \frac{T_1+T_4+10+0}{4} \\ T_3 & = & \frac{T_1+T_4+10+20}{4} \\ T_4 & = & \frac{T_2+T_3+10+20}{4} \end{array}$$

Operando algebraicamente en cada igualdad, podemos escribir las ecuaciones anteriores como

$$\begin{array}{rcl} 4T_1 - T_2 - T_3 & = & 10 \\ -T_1 + 4T_2 - T_4 & = & 10 \\ -T_1 + 4T_3 - T_4 & = & 30 \\ -T_2 - T_3 + 4T_4 & = & 30 \end{array}$$

Es decir, obtenemos un sistema de 4 ecuaciones lineales, con 4 incógnitas, T_1, T_2, T_3, T_4. Este sistema posee solución única.

4) **Problema altimétrico en topografía**

La Topografía es el estudio dimensional de pequeñas porciones de la superficie terrestre. Se estudian básicamente distancias lineales entre puntos definidos. Una distancia que interesa es la distancia vertical entre estos puntos. En cada punto de la Tierra mediante una plomada es posible definir una dirección que se llama Vertical del Lugar. Esta vertical puede materializarse mediante distintos instrumentos, muchos de uso cotidiano. Desde plomadas de albañil hasta los instrumentos topográficos más sofisticados. La vertical permite definir sobre ella un sistema de coordenadas de una dimensión.

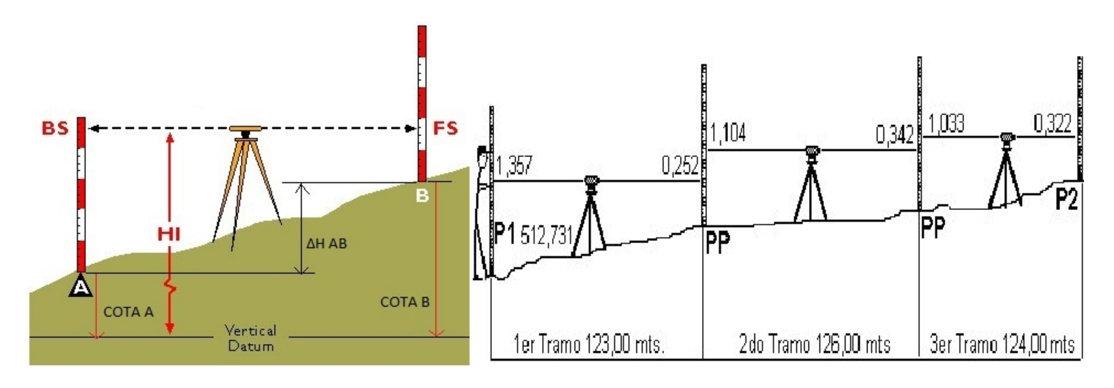


Figura 1.5: Esquema para cálculo de una red de alturas.

Casos particulares son los problemas altimétricos que buscan mediante diversos métodos y procedimientos determinar y representar la *altura* o *cota* de cada punto respecto de un plano de referencia. Con la Altimetría se consigue representar el relieve del terreno mediante planos de curvas de nivel, perfiles, etc.

Para el cálculo de una red de alturas, se *modelan* las *observaciones* para la determinación de las *cotas* (alturas) $x_l, ..., x_n$, donde n especifica la cantidad de puntos. Luego, se miden los *desniveles* o diferencia de alturas, desde el punto i hasta el punto j para dar un valor ΔH_{ij} (probablemente no exacto):

$$\text{Punto i: } x_j - x_i = \Delta H_{ij}$$

Para una red con 3 puntos y 3 mediciones, se obtiene el siguiente sistema de ecuaciones lineales:

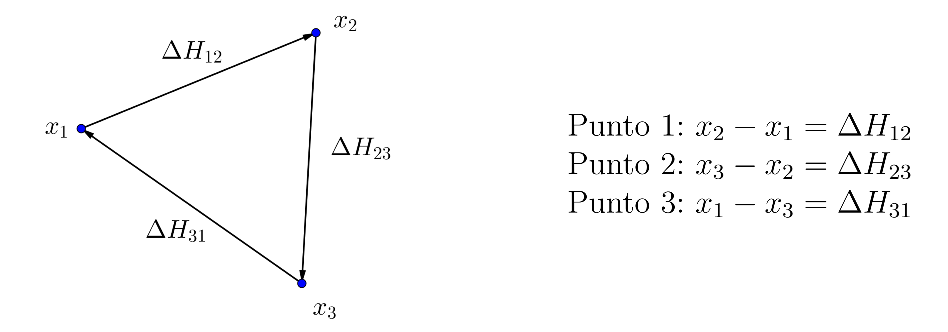


Figura 1.6: Red elemental de tres puntoss.

el cual resulta compatible indeterminado si $\Delta H_{12} + \Delta H_{23} + \Delta H_{31} = 0$ e incompatible en caso contrario.

Los problemas presentados son simples pero pueden ser extendidos a situaciones mucho más complejas, con numerosas (decenas, cientos, miles o más) ecuaciones e incógnitas. Su resolución sistemática requiere del estudio de los conceptos que veremos a continuación.

1.2. Sistemas lineales. Conjunto solución

El caso *más simple* de un sistema de ecuaciones lineales, es el que posee *una sola ecuación lineal* y una sola incógnita x, con a y b constantes reales:

$$ax = b \tag{1.4}$$

Seguramente el lector conoce la *solución* de esta ecuación en caso de que exista:

I. Si a tiene inverso multiplicativo ($a \neq 0$) $\Rightarrow$ la ecuación lineal tiene solución única:

$$x = a^{-1}b$$

(es decir, $x = b/a$) para cualquier valor de b.

II. Si a no tiene inverso multiplicativo ($a = 0$) $\Rightarrow$ la existencia de la solución depende del valor de b:

II.1 Si $b = 0 \Rightarrow$ la ecuación tiene infinitas soluciones (cualquier $x \in \mathbb{R}$ es solución).

II.2 Si $b \neq 0 \Rightarrow$ la ecuación no tiene solución.

Pasemos ahora al caso general. Una ecuación lineal con n incógnitas, $x_1, x_2, \ldots, x_n$, es una ecuación de la forma

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \tag{1.5}$$

donde los coeficientes $a_1, a_2, \ldots, a_n$ y el término b son números reales (o en general complejos) conocidos. Utilizando el símbolo de sumatoria puede escribirse la ecuación (1.5) como

$$\sum_{j=1}^{n} a_jx_j = b \tag{1.6}$$

Si en lugar de una, se tienen varias ecuaciones del tipo anterior en las mismas incógnitas, se obtiene un sistema de ecuaciones lineales:

Definición.

Un sistema de m ecuaciones lineales con n incógnitas $x_1, x_2, \ldots, x_n$, es un sistema de m ecuaciones de la forma

$$\left\{\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\ \vdots & \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m \end{array}\right. \tag{1.7}$$

donde los coeficientes $a_{11}, a_{12}, \ldots, a_{mn}$ y $b_1, \ldots, b_m$ son números reales (o en general complejos) conocidos. El sistema puede escribirse también en forma compacta como

$$\sum_{j=1}^{n} a_{ij}x_j = b_i\,, \qquad i = 1, \ldots, m$$

Según sea el orden del sistema lineal, estos se clasifican en **sistema cuadrado** si $m = n$, es decir, si tiene el mismo número de ecuaciones que de incógnitas, y **sistema rectangular** si $m \neq n$. En este último caso, si el sistema tiene más ecuaciones que incógnitas ($m > n$) se denomina **sobredeterminado**. Si por el contrario tiene menos ecuaciones que incógnitas ($m < n$) se denomina **subdeterminado**:

$$\text{Sistema cuadrado}: \quad m = n \quad (\text{N}^o \text{ de ecuaciones} = \text{N}^o \text{ de incógnitas})$$

$$\text{Sistema rectangular}: \quad m \neq n \quad \begin{cases} \text{Sistema subdeterminado}: & m < n \\ \text{Sistema sobredeterminado}: & m > n \end{cases}$$

Definición.

Una *solución* de un sistema de m ecuaciones con n incógnitas es una n-upla $(x_1, x_2, \ldots, x_n)$ que satisface las m ecuaciones del sistema.

Ejemplo 1.2.1

(a)	(b)	(c)
$\begin{aligned} x_1 + 2x_2 &= 5 \\ 2x_1 + 3x_2 &= 8 \end{aligned}$	$\begin{aligned} x_1 - x_2 + x_3 &= 2 \\ 2x_1 + x_2 - x_3 &= 4 \end{aligned}$	$\begin{aligned} x_1 + x_2 &= 2 \\ x_1 - x_2 &= 4 \\ x_1 + 2x_2 &= 0 \end{aligned}$
(2×2)	(2×3)	(3×2)

Es fácil verificar que:

- En (a) (sistema cuadrado) el par ordenado $(x_1, x_2) = (1, 2)$ satisface ambas ecuaciones, por lo tanto es solución. Se puede verificar también que es la *única solución.*

- En (b) (sistema subdeterminado) la terna $(x_1, x_2, x_3) = (2, 0, 0)$ satisface ambas ecuaciones. Pero también la terna $(x_1, x_2, x_3) = (2, \alpha, \alpha)$ donde α es un número real cualquiera, satisface ambas ecuaciones. En este caso, existen pues *infinitas soluciones* porque hay infinitas ternas (3-uplas) que satisfacen el sistema.

- En (c) (sistema sobredeterminado) *no existe solución*: Si sumamos las dos primeras ecuaciones obtenemos $2x_1 = 6$, de donde $x_1 = 3$. Utilizando ahora la primera o la segunda ecuación, se obtiene $x_2 = -1$. Pero estos valores implican $x_1 + 2x_2 = 1$, lo que está en contradicción con la última ecuación. Por lo tanto, este sistema no tiene un par (x_1, x_2) que satisfaga estas tres ecuaciones a la vez.

Debemos remarcar, no obstante, que no todo sistema cuadrado es compatible o posee solución única, que no todo sistema subdeterminado es compatible, y que no todo sistema sobredeterminado es incompatible, como muestran los siguientes ejemplos:

Ejemplo 1.2.2

(a_1)	(a_2)	(b)	(c)
$x_1 + 2x_2 = 5$ $2x_1 + 4x_2 = 10$ (2×2)	$x_1 + 2x_2 = 5$ $2x_1 + 4x_2 = 0$ (2×2)	$x_1 - x_2 + x_3 = 2$ $x_1 - x_2 + x_3 = 3$ (2×3)	$x_1 + x_2 = 2$ $x_1 - x_2 = 4$ $x_1 + 2x_2 = 1$ (3×2)

Es fácil verificar que:

- En (a_1) (sistema cuadrado) la segunda ecuación es la primera multiplicada por dos, por lo que no aporta una nueva condición. Es fácil verificar entonces que todo par de la forma $(x_1, x_2) = (5 - 2\alpha, \alpha)$ es solución del sistema para cualquier α real, por lo que el sistema posee infinitas soluciones.

- En (a_2) (sistema cuadrado) es claro que si $x_1 + 2x_2 = 5$, entonces $2x_1 + 4x_2 = 2(x_1 + 2x_2) = 10$ no puede ser igual a 0. Este sistema es entonces *incompatible*.

- En (b) (sistema subdeterminado) es evidente que las dos ecuaciones son incompatibles, pues si $x_1 - x_2 + x_3$ es 2, no puede ser a la vez 3. Este sistema es entonces incompatible, a pesar de ser subdeterminado (más incógnitas que ecuaciones).

- En (c) (sistema sobredeterminado) vimos en el ejemplo anterior que las dos primeras ecuaciones implican $x_1 = 3$, $x_2 = -1$. Y estos valores ahora sí satisfacen la tercera ecuación. Por lo tanto, este sistema es compatible con solución única $(x_1, x_2) = (3, -1)$, a pesar de que es sobredeterminado (más ecuaciones que incógnitas).

Demostraremos luego que al igual que en el caso (1.4) de una ecuación lineal con una incógnita, y tal como vimos en estos ejemplos, todo sistema de ecuaciones lineales puede o bien tener solución única, o bien tener infinitas soluciones o no tener ninguna solución:

Definición.

Un sistema con al menos una solución se denomina **sistema compatible o consistente**.

Si la **solución es única**, se lo denomina **sistema compatible determinado**.

Si existen **infinitas soluciones** se lo llama sistema **compatible indeterminado**.

Un sistema **sin solución** se llama **sistema incompatible o inconsistente**.

Al conjunto de todas las soluciones de un sistema se lo llama **conjunto solución**.

$$\text{Sistema:} \begin{cases} \text{Compatible:} \begin{cases} \text{Determinado} \\ \text{Indeterminado} \end{cases} \\ \text{Incompatible} \end{cases}$$

1.2.1. Interpretación geométrica

Una ecuación lineal $ax + by = c$, con dos incógnitas x y y, es posible interpretarla geométricamente como la ecuación cartesiana de una *recta* en $\mathbb{R}^2$. Entonces, resolver un sistema de dos ecuaciones ($m = 2$) con dos incógnitas ($n = 2$), es decir, encontrar pares (x, y) que satisfagan ambas ecuaciones, es equivalente a analizar si dos rectas en el plano se intersecan en un punto, si son coincidentes, o si son paralelas. Por ejemplo:

I.

$$\begin{cases} x + y &= 2 \\ x + 2y &= 1 \end{cases}$$

Solución única: $(x, y) = (3, -1)$. Sistema compatible determinado.

Geométricamente, las ecuaciones corresponden a rectas no paralelas. La solución única $(x, y) = (3, -1)$ es el punto donde se cortan.

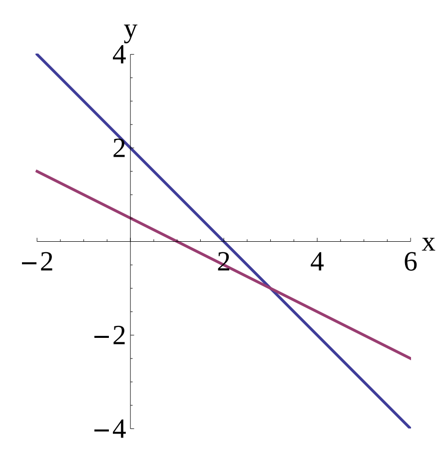


Figura 1.7: Sistema de 2×2 compatible determinado.

II.

$$\begin{cases} x + y &= 2 \\ x + y &= 0 \end{cases}$$

Sin solución. Sistema de 2×2 incompatible.

Geométricamente, las ecuaciones corresponden a rectas paralelas no coincidentes. No tienen puntos en común.

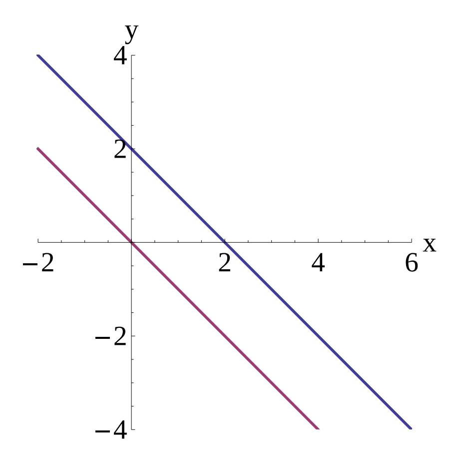


Figura 1.8: Sistema de 2×2 incompatible.

III.

$$\begin{cases} x+y &= 2 \\ 2x+2y &= 4 \end{cases}$$

Infinitas soluciones: $(x, y) = (2-\alpha, \alpha)$, $\alpha \in \mathbb{R}$. Sistema compatible indeterminado.

Geométricamente, las ecuaciones corresponden a rectas coincidentes. El conjunto solución $\{(2-\alpha, \alpha), \alpha \in \mathbb{R}\}$ es el conjunto de puntos de esta recta.

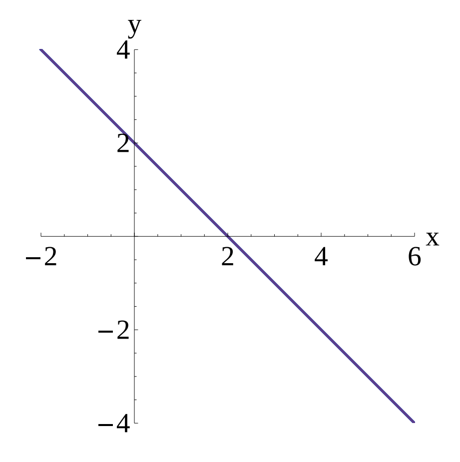


Figura 1.9: Sistema de 2×2 compatible indetederminado.

IV.

$$\begin{cases} x+y &= 2 \\ x-y &= 4 \\ x+2y &= 0 \end{cases}$$

Sistema incompatible.

Geométricamente, las ecuaciones corresponden a tres rectas no paralelas, que no se cruzan todas en un mismo punto.

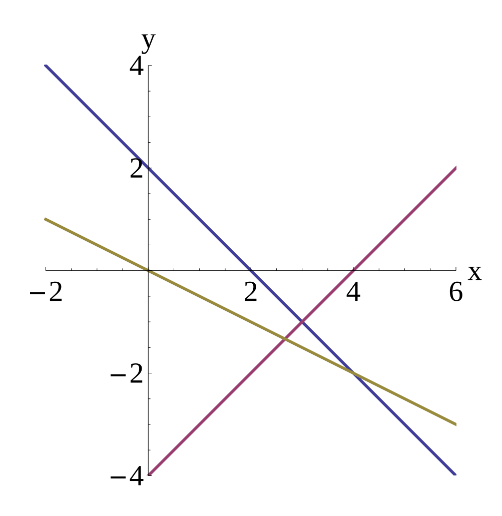


Figura 1.10: Sistema de 3×2 incompatible.

V.

$$\begin{cases} x+y &= 2 \\ x-y &= 4 \\ x+2y &= 1 \end{cases}$$

Sistema compatible determinado: $(x, y) = (3, -1)$.

Geométricamente, las ecuaciones corresponden a tres rectas no paralelas, que se intersecan todas en un mismo punto.

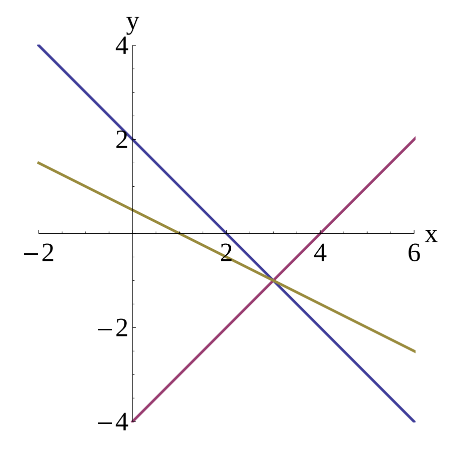


Figura 1.11: Sistema de 3×2 compatible determinado.

Problema 1.2.1
La ecuación cartesiana de un *plano* $ax + by + cz = d$ en $\mathbb{R}^3$, es algebraicamente una ecuación lineal con tres incógnitas. Analizar en términos geométricos, como se hizo en el caso de rectas en el plano, los posibles tipos de conjunto solución que pueden ocurrir con sistemas de 2 y 3 ecuaciones lineales con 3 incógnitas x, y, z (sistemas 2×3 y 3×3). Algunos de los posibles casos son mostrados en la figura siguiente.

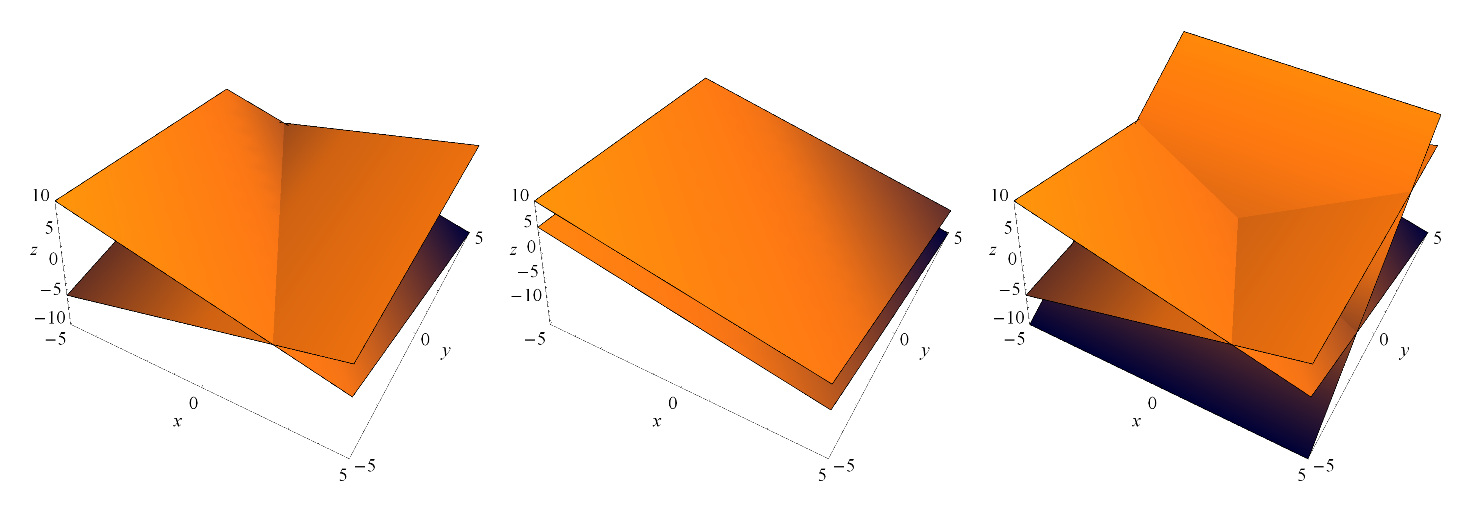


Figura 1.12: Representación geométrica de sistemas lineales con tres incógnitas ($m \times 3$). Izquierda: Sistema de 2×3 (dos ecuaciones) compatible indeterminado. La intersección de dos planos no paralelos es una recta. Centro: Sistema de $2{\times}3$ incompatible (planos paralelos no coincidentes). Derecha: Sistema de 3×3 (tres ecuaciones) compatible determinado. La intersecci´on de tres planos no paralelos es un punto.

1.2.2. Sistemas homogéneos

Definición.
En el caso que todas las constantes b_i en (1.7) sean cero, el sistema de ecuaciones lineales se denomina ***homogéneo***:

$$\left\{\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & 0 \\ \vdots & \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & 0 \end{array}\right. \tag{1.8}$$

o sea, $\sum_{j=1}^{n} a_{ij}x_j = 0$, $i = 1, \ldots, m$.

Es obvio que estos sistemas siempre poseen al menos la *solución trivial o nula*

$$x_1 = x_2 = \ldots = x_n = 0$$

Por lo tanto, un sistema homogéneo **siempre es compatible**. Puede ser:

I. Compatible determinado (la solución trivial es la única solución)

II. Compatible indeterminado (solución trivial + infinitas soluciones no triviales)

$$\text{Sistema homogéneo:} \left\{\begin{array}{l} \text{Compatible determinado} \\ \text{Compatible indeterminado} \end{array}\right.$$

Ejemplo 1.2.3

Dado el siguiente siguiente sistema homogéneo,

$$\left\{\begin{array}{r} x_1 + 2x_2 = 0 \\ -x_1 - 2x_2 = 0 \end{array}\right.$$

es fácil verificar que el par ordenado $(x_1, x_2) = (0, 0)$ es solución del sistema (solución trivial). Pero también es solución cualquier par de la forma $(-2\alpha, \alpha)$ con α un número real cualquiera. Es decir, el sistema posee *infinitas soluciones*, y no sólo la solución trivial, siendo entonces compatible indeterminado.

En cambio, en el siguiente sistema homogéneo,

$$\left\{\begin{array}{r} 3x + 2y + z = 0 \\ y + z = 0 \\ -2z = 0 \end{array}\right.$$

es fácil verificar que tiene la solución única $(x, y, z) = (0, 0, 0)$. Este sistema es entonces compatible determinado.

1.3. Sistemas equivalentes

Definición.
Dos (o más) sistemas lineales con el mismo conjunto de variables o incógnitas se dicen **equivalentes** sí y sólo sí tienen el mismo *conjunto solución*.

Ejemplo 1.3.1

(a)	(b)
$3x_1 + 2x_2 - x_3 = -2$	$3x_1 + 2x_2 - x_3 = -2$
$x_2 = 3$	$-3x_1 - x_2 + x_3 = 5$
$2x_3 = 4$	$3x_1 + 2x_2 + x_3 = 2$

Estos dos sistemas de ecuaciones, son *equivalentes*. Ambos tienen 3 incógnitas y el mismo conjunto solución: $(x_1, x_2, x_3) = (-2, 3, 2)$. Observar además, que la primera ecuación de ambos sistemas es la misma. Mientras que en (a) las restamtes ecuaciones dicen que $x_2 = 3$ y $x_3 = 2$, en (b), si sumamos la primera ecuación a la segunda se obtiene $x_2 = 3$ y si restamos la primera ecuación a la tercera se obtiene $2x_3 = 4$, o sea, $x_3 = 2$.

Por otro lado, cualquier solución del sistema (a) debe ser también solución del sistema (b), porque restando en (a) la primera ecuación a la segunda, se obitene la segunda ecuación del sistema (b), y sumando la primera y tercera ecuación del sistema (a), se obtiene la tercera ecuación del sistema (b). Es decir, que realizando *operaciones algebraicas* sobre las ecuaciones de un sistema, es posible "pasar al otro".

1.3.1. Operaciones elementales

Definición.
Llamaremos *operaciones elementales* a las operaciones algebraicas sobre las ecuaciones de un sistema lineal que no modifican *el conjunto solución*. Esto quiere decir, que la aplicación de tales operaciones producen *sistemas* $m \times n$ *equivalentes*.
Estas operaciones son:

1. Cambiar el orden de dos ecuaciones (permutar dos ecuaciones).
2. Multiplicar una o más ecuaciones por una constante *distinta de cero* (cambio de escala de los coeficientes de las ecuaciones).
3. Sumar (o restar) a una ecuación particular el múltiplo de otra ecuación del sistema.

Observación. Multiplicar una ecuación por 0 **no** está permitido, ya que esto puede cambiar el conjunto solución (¿por qué?). Y sumar a una ecuación un multiplo de si misma es obviamente equivalente a multiplicarla por una constante (¡justificar!).

En lo que sigue, utilizaremos estas operaciones elementales para obtener **sistemas equivalentes** más fáciles de resolver, tales como los **sistemas triangulares**.

1.3.2. Sistema triangular

Un sistema 3×3 de la forma

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{33}x_3 &= b_3 \end{aligned}$$

con a_{33}, a_{22}, a_{11} no nulos, es **fácil de resolver** por *sustitución* (comenzando desde abajo hacia arriba):

$$\begin{aligned} x_3 &= \frac{b_3}{a_{33}} \\ x_2 &= \frac{b_2 - a_{23}x_3}{a_{22}} \\ x_1 &= \frac{b_1 - a_{12}x_2 - a_{13}x_3}{a_{11}} \end{aligned}$$

En este caso especial "triangular" la solución del sistema $n \times n$ **es única**, pués a_{33}, a_{22}, a_{11} son no nulos. A esta forma la denominaremos *sistema triangular*.

Definición.
Se dice que un sistema cuadrado de orden $n \times n$ es de **forma triangular** si para cada ecuación k-ésima, $k = 1, \ldots, n$, los coeficientes de sus primeras $(k-1)$ variables son "cero", y el coeficiente de la variable x_k es distinto de cero. Es decir, su forma es

$$a_{kk}x_k + a_{kk+1}x_{k+1} + \cdots + a_{kn}x_n = b_k \qquad \text{con } a_{kk} \neq 0, \quad k = 1, \ldots, n\,.$$

Entonces en general, para analizar y resolver un sistema de ecuaciones lineales del tipo $n \times n$, realizaremos operaciones elementales para generar, en caso de ser posible, un "sistema equivalente" en forma triangular.

Ejemplo 1.3.1 Para resolver el sistema

$$\begin{aligned} 3x_3 &= 9 \\ x_1 + 5x_2 - 2x_3 &= 2 \\ \tfrac{1}{3}x_1 + 2x_2 \qquad &= 3 \end{aligned}$$

lo transformamos usando repetidamente operaciones elementales, hasta que tenga una forma fácil de resolver y si es posible triangular:

$$\xrightarrow[(f_1 \leftrightarrow f_3)]{\text{se permuta la ecuación (fila) 1 con la ecuación 3}} \quad \begin{aligned} \tfrac{1}{3}x_1 + 2x_2 \qquad &= 3 \\ x_1 + 5x_2 - 2x_3 &= 2 \\ 3x_3 &= 9 \end{aligned}$$

$$\xrightarrow[(3f_1)]{\text{se multiplica la ecuación 1 por 3}} \quad \begin{aligned} x_1 + 6x_2 \qquad &= 9 \\ x_1 + 5x_2 - 2x_3 &= 2 \\ 3x_3 &= 9 \end{aligned}$$

$$\xrightarrow[(f_2 - f_1)]{\text{se suma a la ecuación 2 la ecuación 1 multiplicada por -1}} \quad \begin{aligned} x_1 + 6x_2 \qquad &= 9 \\ -x_2 - 2x_3 &= -7 \\ 3x_3 &= 9 \end{aligned}$$

El único paso no trivial es el realizado en tercer término. Hemos multiplicado ambos miembros de la primer ecuación por -1, y sumado ese resultado a la segunda ecuación, para luego escribir ese resultado como la nueva segunda ecuación, reemplazando a la original. Ahora, se puede encontrar el valor de cada variable fácilmente. En otros casos, puede ocurrir que no se obtenga una forma triangular.

Veamos algunos ejemplos simples, indicando cada operación realizada para pasar de un sistema a otro equivalente.

Ejemplo 1.3.2 Consideremos el siguiente sistema de 2×2

$$\begin{aligned} x+2y&=8 \\ 2x+4y&=8 \end{aligned} \quad \underset{f_2-2f_1}{\overset{\text{se suma a la ecuación 2 la ecuación 1 multiplicada por } -2}{\longrightarrow}} \quad \begin{aligned} x+2y&=\ \ 8 \\ 0&=-8 \end{aligned}$$

En este caso se advierte que el sistema equivalente es *incompatible* (hay una ecuación inconsistente).

Ejemplo 1.3.3 En el sistema

$$\begin{aligned} x+\ \ y&=4 \\ 2x+2y&=8 \end{aligned}$$

es evidente que cualquier par x, y de números que satisface la primer ecuación también satisface la segunda.

La solución, es el conjunto de pares: $\{(x,y) \mid x+y=4\}$. Algunas soluciones son: $(0,4)$, $(-1,5)$, y $(10,-6)$. Si se hubiesen aplicado operaciones elementales para intentar llevarlo a la forma triangular, se obtendría

$$\underset{f_2-2f_1}{\overset{\text{se suma a la ecuación 2 la ecuación 1 multiplicada por -2}}{\longrightarrow}} \quad \begin{aligned} x+y&=4 \\ 0&=0 \end{aligned}$$

En este caso el sistema lineal tiene infinitas soluciones. Es *compatible indeterminado.*

Comentario. La igualdad que aparece en este ejemplo: "$0 = 0$" es un "indicador" de que la segunda ecuación es "redundante" (no aporta nueva información). Por ser un sistema de 2×2, eso ya es suficiente para saber que existen infinitas soluciones. En general, en sistemas más grandes, la expresión "$0 = 0$" no es suficiente para derivar esa conclusión.

Problema 1.3.1
a) Resolver el siguiente sistema, aplicando operaciones elementales para llegar a una forma triangular:

$$\left\{ \begin{array}{rcl} x_1+2x_2+x_3 &=& 3 \\ 3x_1-x_2-3x_3 &=& -1 \\ 2x_1+3x_2+x_3 &=& 4 \end{array} \right.$$

Verificar que por sustitución hacia atrás resulta: $(x_1, x_2, x_3) = (3, -2, 4)$.
b) Resolver el sistema homogéneo asociado ¿Es necesario para esto, realizar cálculos extras a los realizados en la parte *a)*?

Problema 1.3.2
Aplicar operaciones elementales con el objetivo de llevar, de ser posible, el siguiente sistema a una forma triangular. Decidir cuántas soluciones tiene.

$$
\begin{aligned}
x + \ y + \ z &= \ 4 \\
2x + 2y + 2z &= \ 8 \\
4x + 4y + 4z &= 20
\end{aligned}
$$

Problema 1.3.3

Analizar si es posiblle llevar el sistema siguiente de 4×4 a la forma triangular. ¿Cuántas soluciones tiene el sistema? ¿Por qué? En el sistema equivalente, ¿Existe algún sistema o subsistema triangular respecto de algunas de las variables?

$$
\begin{aligned}
x_1 + 6x_2 \qquad\quad + \ x_4 &= \ 9 \\
- \ x_2 - 2x_3 + \ x_4 &= -7 \\
3x_3 + \ x_4 &= \ 9 \\
- \ x_2 + \ x_3 + 2x_4 &= \ 2
\end{aligned}
$$

Problema 1.3.4

Analizar para distintos valores de k y en caso de ser compatible, dar el conjunto solución. Realizar el análisis de dos formas distintas. (a) Geométricamente: utilizando un software, graficar las dos rectas que dependen del parámetro k, y realizar el análisis geométrico. (b) Algebraicamente: llevarlo a un sistema equivalente mediante operaciones elementales.

$$
\begin{aligned}
kx + \ y &= 1 \\
x + ky &= 1
\end{aligned}
$$

1.4. Matriz de coeficientes

1.4.1. Matriz de coeficientes de un sistema. Matriz ampliada

Para simplificar la descripción de los cálculos mediante operaciones elementales en los procedimientos previamente descriptos, conviene introducir el concepto de *matriz.* Básicamente una matriz es una tabla de doble entrada (filas y columnas) en la que es posible disponer objetos matemáticos. En este caso extraeremos los coeficientes del sistema lineal de ecuaciones y los dispondremos en una tabla.

Definición.
La **matriz de coeficientes** A de un sistema $m \times n$ es:

$$
A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \tag{1.9}
$$

y la **matriz ampliada** $(A \mid b)$ del mismo sistema $m \times n$ es:

$$
(A \mid b) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right) \tag{1.10}
$$

Realizar las operaciones elementales sobre las ecuaciones lineales es equivalente a realizar las siguientes operaciones sobre las filas de la matriz ampliada:

1. Intercambiar dos filas.

2. Multiplicar una o más filas por un número no nulo.

3. Sumar a una fila j-ésima un múltiplo de otra fila i-ésima, y reemplazar la fila j por "la resultante de la operación realizada".

Notaciones.
Cuando explicitamos las operaciones realizadas para reducir el sistema por este método, abreviamos 'fila i' mediante f_i.
La operación que permuta la fila i con la fila j $(i \neq j)$ se denotará $f_i \leftrightarrow f_j$.
La operación que multiplica la fila i por un número $\alpha \neq 0$ se denotará αf_i.
La operación que suma a la fila j el resultado de multiplicar por un número α la fila i se denotará $f_j + \alpha f_i$.
También, para ahorrar escritura, se listarán los pasos del último tipo juntos, cuando se use la misma fila f_i.

1.4.2. Pivoteo

¿Cómo realizar las operaciones elementales en forma sistemática?
Para cada fila "no nula", denominaremos **pivote** al primer elemento no nulo de esa fila.

Apoyados en el "pivote", mediante operaciones elementales, se llevan a " 0 " todos los términos de esa columna que están por debajo del "pivote" de acuerdo al siguiente procedimiento:

1. Tomar la primer fila y su "primer coeficiente como pivote". Con operaciones adecuadas eliminar los primeros coeficientes de las filas siguientes (o sea, los elementos de la primer columna, salvo el pivote de la fila 1):

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 3 & -1 & -3 & -1 \\ 2 & 3 & 1 & 4 \end{array}\right) \underset{f_3-2f_1}{\overset{f_2-3f_1}{\longrightarrow}} \left(\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & -7 & -6 & -10 \\ 0 & -1 & -1 & -2 \end{array}\right)$$

2. Luego, considerar la segunda fila y su segundo coeficiente como "pivote":

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & -7 & -6 & -10 \\ 0 & -1 & -1 & -2 \end{array}\right) \overset{f_3-1/7f_2}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & -7 & -6 & -10 \\ 0 & 0 & -1/7 & -4/7 \end{array}\right)$$

3. Continuar así, con la tercer fila y el elemento de la tercer columna, hasta obtener una forma triangular (no necesariamente única, ¿porqué?).

   Si durante este proceso, el coeficiente que correspondería ser pivote de una fila particular resulta ser cero, entonces se permuta esa fila con alguna de las que le siguen, para lograr una fila con un pivote no nulo. Si no existe una fila (entre las que le siguen) que tenga un coeficiente no nulo en esa columna que se analiza, se abandona esa columna y se pasa a la columna siguiente en la misma fila.

4. Finalmente, resolver *por sustitución* hacia atrás (como antes).

Observación. Veamos un ejemplo del caso que acabamos de explicar, es decir cuando, en la columna que se está analizando, no hay ninguna posibilidad de obtener un pivote no nulo al permutar con las las filas que le siguen. ¿Qué se hace en este caso?

Ejemplo 1.4.1 Sistema 5×5

$$\left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 0 & 0 & 1 & -1 \\ -2 & -2 & 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 1 & 3 & -1 \\ 1 & 1 & 2 & 2 & 4 & 1 \end{array}\right) \underset{\substack{f_2+f_1 \\ f_3+2f_1 \\ f_5-f_1}}{\longrightarrow} \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 2 & 2 & 5 & 3 \\ 0 & 0 & 1 & 1 & 3 & -1 \\ 0 & 0 & 1 & 1 & 3 & 0 \end{array}\right)$$

Todos los posibles pivotes en la columna 2 son ceros. Entonces debemos tomar un pivote en la misma fila pero en la columna 3 y continuar el proceso:

$$\left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 2 & 2 & 5 & 3 \\ 0 & 0 & 1 & 1 & 3 & -1 \\ 0 & 0 & 1 & 1 & 3 & 0 \end{array}\right) \underset{\substack{f_3-2f_2 \\ f_4-f_2 \\ f_5-f_2}}{\longrightarrow} \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right)$$

De nuevo, las posibles elecciones del pivote en la columna 4 son ceros. Entonces nos movemos a la columna 5 en esa misma fila:

$$\left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right) \underset{\substack{f_4-f_3 \\ f_5-f_3}}{\longrightarrow} \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 & 0 & -3 \end{array}\right)$$

Llegamos así a una matriz *de forma escalonada*. Las últimas filas representan las ecuaciones

$$\begin{array}{rcl} 0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 & = & -4 \\ 0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 & = & -3 \end{array}$$

Por tanto, se deduce que este sistema es incompatible, no existe solución. Esto es en realidad evidente ya a partir del primer paso (¿por qué?), pero hemos realizado el procedimiento sistemático completo para mostrar la forma de proceder en el caso general.

Ejemplo 1.4.2 Si ahora realizamos las mismas operaciones elementales sobre el sistema que ha cambiado, respecto del previo, sólo los términos b_i de la derecha, obtenemos:

$$\left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 0 & 0 & 1 & -1 \\ -2 & -2 & 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 1 & 3 & 3 \\ 1 & 1 & 2 & 2 & 4 & 4 \end{array}\right) \longrightarrow \cdots \longrightarrow \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Ahora, las dos últimas filas representan la ecuación $0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 = 0$ que es satisfecha por cualquier 5-upla $(x_1, x_2, x_3, x_4, x_5)$. Por tanto, en este nuevo sistema el conjunto solución son todas las 5-uplas que satisfacen

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 + x_5 &= 1 \\ x_3 + x_4 + 2x_5 &= 0 \\ x_5 &= 3 \end{aligned}$$

Dentro de estas tres ecuaciones, notamos que tenemos dos tipos de variables: *variables independientes (o libres) y variables dependientes*

$$\begin{aligned} x_1, x_3, x_5 &= \text{variables dependientes:} \\ x_2, x_4 &= \text{variables independientes} \end{aligned}$$

Moviendo las variables independientes al término de la derecha

$$\begin{aligned} x_1 + x_3 + x_5 &= 1 - x_2 - x_4 \\ x_3 + 2x_5 &= -x_4 \\ x_5 &= 3 \end{aligned}$$

se obtiene un sub sistema triangular, respecto de las variables x_1, x_3, x_5, que son las variables dependientes. Por lo tanto, estas tres variables se pueden despejar en función del par de valores (α, β) asignados a (x_2, x_4). El sistema triangular de las "variables dependientes" tiene solución única para cada par de valores (α, β). Así,

$$\begin{aligned} x_5 &= 3 \\ x_3 &= -x_4 - 2x_5 = -\beta - 6 \\ x_1 &= (1 - x_2 - x_4) - (x_3 + x_5) = 4 - \alpha \end{aligned}$$

El conjunto solución del sistema dado, resulta:

$$(x_1, x_2, x_3, x_4, x_5) = (4 - \alpha, \alpha, -\beta - 6, \beta, 3)$$

donde α y β son números reales cualesquiera. En este caso se encuentra que el sistema tiene "infinitas soluciones" porque el sistema original de 5×5 resultó ser equivalente a un sistema de 3×5. Dos de las cinco ecuaciones originales son *redundantes* y no agregan nueva información sobre las incógnitas $(x_1, x_2, x_3, x_4, x_5)$.

Problema 1.4.1

En el ejemplo anterior, mantener la "misma matriz" de coeficientes pero cambiar la columna de la derecha de la matriz ampliada (los b_i) por ceros. Es decir, considerar el sistema homogéneo asociado. Analizar que tipo de solución tiene y obtener el conjunto solución.

Importante. Como vimos en los dos ejemplos anteriores y en el ejercicio previo, el hecho que un sistema "no tenga solución" o tenga "infinitas soluciones" depende de las constantes $\{b_1, b_2, b_3, b_4, b_5\}$. Nos preguntamos entonces si existe algúna propiedad o característica de los coeficientes de la matriz del sistema $n \times n$, que pueda decirnos cuando el sistema tiene una única solución, y cuando no tiene solución o tiene infinitas soluciones, sin tener necesidad de resolver efectivamente el sistema. Más adelante, veremos que analizando la matriz de coeficientes del sistema y la matriz ampliada será posible saber la respuesta.

1.5. Método de eliminación de Gauss

Un algoritmo eficiente y fácilmente programable para resolver sistemas lineales de cualquier tamaño con la metodología ya introducida es el Método de Gauss. Comenzamos definiendo una forma especial de matrices.

Definición.
Una matriz M es de forma *escalonada por filas* si:

1. Cualquier fila nula (es decir, que tenga sólo ceros) está por debajo de las filas que tienen algún elemento no nulo.

2. Si una fila j, $j > 1$, es no nula, entonces el número de ceros previos al primer elemento no nulo (pivote) debe ser estrictamente mayor que el número de ceros previos al pivote de la fila anterior (la fila $j-1$).

Por lo tanto, un sistema está en la *forma escalonada por filas* si en cada fila la variable pivote está a la derecha de la variable pivote de la fila previa a ella.

Observación 1. En algunos textos se pide también que el primer coeficiente no nulo (pivote) de cada fila no nula **sea** 1. Esta condición requiere el paso adicional de multiplicar cada fila no nula por una constante adecuada. Este paso es conveniente para poder despejar en forma directa las variables dependientes, pero no es imprescindible.

Observación 2. Cuando una matriz está en forma escalonada, los primeros elementos diferentes de cero de cada fila, reciben el nombre de **pivotes**. Note que por ser el pivote el primer elemento no nulo de la fila no hay forma que una fila tenga más de un pivote: puede no tener pivote en caso de que sea una fila de ceros (fila nula), pero no puede tener dos o más. Notar también que por estar escalonada la matriz, no hay forma que dos pivotes queden en la misma columna: puede una columna no tener pivote, pero si tiene pivote no puede tener dos o más. De este hecho, se concluye que una matriz de $m \times n$ no puede tener más de m pivotes porque tiene a lo sumo uno por cada fila.

Definición.
El proceso que utiliza operaciones elementales sobre las filas para reducir un sistema lineal cualquiera a un *sistema escalonado por filas*, se denomina **Método de eliminación Gaussiana** o **Método de reducción por filas**.

Los pasos siguientes permiten llevar una matriz, mediante operaciones elementales sistemáticas sobre sus filas, a una matriz en forma escalonada:

1. Disponga en una matriz ampliada los coeficientes del sistema de ecuaciones lineales y del término independiente.

2. Si $a_{11} \neq 0$ tómelo como pivote. Caso contrario permute la fila 1 por una fila que no tenga cero en el elemento de la primera columna.

3. Mediante operaciones elementales sobre las filas de la matriz resultante, obtenga ceros por debajo del elemento pivote. Esto se logra restando a la fila i la fila 1 multiplicada por a_{i1}/a_{11}, es decir, mediante las operaciones $f_i - \frac{a_{i1}}{a_{11}} f_1$ para $i \geq 2$, tal que $a_{i1} \to 0$ y $a_{ij} \to a_{ij} - \frac{a_{i1}}{a_{11}} a_{1j}$ para $i \geq 2$ y $j \geq 2$.

4. Si el elemento a_{22} de la matriz resultante es no nulo, tómelo como pivote. En caso contrario permute esta fila por otra debajo de ella que no tenga cero en el elemento de la segunda columna. Si no hubiese ninguna fila por debajo con un elemento no nulo en la columna 2, pase a la columna siguiente y repita el procedimiento anterior, hasta obtener un elemento pivote.

5. Repita los pasos 3 y 4 para las filas siguientes a partir del nuevo pivote. Continúe el proceso hasta llegar a la última fila no nula. Se obtiene como resultado final una matriz en forma escalonada.

A partir de la forma escalonada se obtienen las siguientes conclusiones, válidas para sistemas generales $m \times n$:

Conclusión 1.
Si la forma escalonada por filas de la matriz ampliada **incluye** alguna fila de la forma

$$\left(\begin{array}{cccc|c} 0 & 0 & \cdots & 0 & b \end{array}\right), \qquad b \neq 0,$$

entonces el sistema es **incompatible** (sin solución).
En efecto, tal fila implica $0x_1 + 0x_2 + \ldots + 0x_n = b \neq 0$, lo cual no puede ser satisfecho por ninguna n-upla $(x_1, x_2, \ldots, x_n)$.

Conclusión 2.
Si no existe ninguna fila como la indicada en la conclusión 1, el sistema es **compatible**. Existen entonces dos posibilidades:

1. Si las filas **no nulas** de la forma escalonada por filas de la matriz ampliada, forman un sistema triangular (respecto de todas las variables del sistema), es decir, del tipo

$$\left(\begin{array}{cccc|c} 1 & \times & \times & \times & \times \\ 0 & 1 & \times & \times & \times \\ 0 & 0 & 1 & \times & \times \\ 0 & 0 & 0 & 1 & \times \end{array}\right)$$

   entonces el sistema tiene solución **única**: sistema **compatible determinado**. Partiendo de la última fila, se obtiene un valor único para todas las incógnitas.

   Este caso **no puede darse en los sistemas subdeterminados** ($m < n$), pues requiere un número total de filas al menos igual al número de columnas.

2. En caso contrario, existe una sub-matriz triangular correspondiente a un subconjunto de las variables (las variables dependientes o "pivotes", que corresponden a las columnas con pivotes), y las restantes son variables **libres** (independientes), que pueden tomar cualquier valor. Por lo tanto, existen **infinitas** soluciones: sistema **compatible indeterminado**.

   Este caso puede ocurrir tanto si $m < n$ como si $m = n$ o $m > n$.

Por ejemplo, si la forma escalonada de la matriz ampliada es de la forma

$$\left(\begin{array}{cccccc|c} 1 & \times & \times & \times & \times & \times & \times \\ 0 & 0 & 1 & \times & \times & \times & \times \\ 0 & 0 & 0 & 1 & \times & \times & \times \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

(donde las $\times$ indican números arbitrarios) el sistema es compatible indeterminado, siendo x_2, x_5 y x_6 las variable libres y x_1, x_3 y x_4 las variables dependientes.

Conclusión 3.
De las conclusiones 1 y 2 anteriores, vemos que:
1. Un sistema subdeterminado ($m < n$) sólo puede ser **compatible indeterminado** o **incompatible**.
2. Los sistemas cuadrados ($m = n$) o sobredeterminados ($m > n$) pueden ser compatibles determinados, compatibles indeterminados o incompatibles, dependiendo del caso.

Conclusión 4.
En relación a los **sistemas homogéneos** (las constantes b_i son cero) podemos concluir:

- Son siempre **compatibles**, porque $(x_1, x_2, \ldots, x_n) = (0, 0, \ldots, 0)$ es siempre una solución (solución trivial).
- Si es compatible determinado, la única solución es la trivial $(x_1, \ldots, x_n) = (0, \ldots, 0)$.
- Si es compatible indeterminado, el sistema posee, además de la solución trivial, infinitas soluciones no triviales $(x_1, x_2, \ldots, x_n)$, con los x_i no todos nulos.
- Un sistema homogéneo **subdeterminado** ($m < n$) es necesariamente **compatible indeterminado** (tiene infinitas soluciones), por la conclusión 3.1

Problema 1.5.1 Analizar lo expuesto previamente para el caso de un sistema cuadrado $n \times n$ homogéneo.

Comentario.
Los sistemas sobredeterminados ($m > n$) no homogéneos, no son necesariamente incompatibles, aunque frecuentemente lo son (¿puede explicar el lector por qué?).
Y los sistemas subdeterminados ($m < n$) no homogéneos pueden ser, como hemos visto, incompatibles, aunque frecuentemente son compatibles indeterminados (¿puede explicar el lector por qué?).

Ejemplo 1.5.2 Consideremos el sistema

$$\begin{aligned} x - y + z &= 1 \\ 3x + z &= 3 \\ 5x - 2y + 3z &= 5 \end{aligned}$$

Reduciéndolo por filas, obtenemos

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 3 & 0 & 1 & 3 \\ 5 & -2 & 3 & 5 \end{array}\right) \underset{f_3-5f_1}{\overset{f_2-3f_1}{\longrightarrow}} \left(\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 0 & 3 & -2 & 0 \\ 0 & 3 & -2 & 0 \end{array}\right) \underset{f_3-f_2}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 0 & 3 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Se observa que es equivalente a un sistema de 2×3, que tiene una matriz triangular de 2×2 para las variables pivotes x e y, mientras que la variable z es libre.

Existen entonces “infinitas soluciones”dependientes de un parámetro z (variable independiente o libre). El conjunto solución es:

$$\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + z \begin{pmatrix} -1/3 \\ 2/3 \\ 1 \end{pmatrix} \mid z \in \mathbb{R} \right\}$$

Notar que la solución es la suma de dos vectores. El primero es una solución del sistema no homogéneo (la obtenida para $z = 0$) mientras que el segundo, dependiente del parámetro libre z, es solución del sistema **homogéneo** asociado (es independiente de los b_i y por ende el único término en el caso homogéneo $b_i = 0 \ \forall \ i$).

Importante
En todo sistema no homogéneo compatible indeterminado, es posible descomponer la solución en la suma de dos partes: una es una solución “particular” del sistema no homogéneo (es la parte que no contiene variables libres) y la otra, la soluci´on del sistema homog´eneo asociado (la parte que contiene las variables libres). Esto ser´a demostrado en los próximos capítulos, pero tal como en el ejemplo previo, puede ya verse en general a partir de la forma escalonada de la matriz ampliada (¡verificar!).

Problema 1.5.3 Resolver el siguiente sistema de 4×3 tratando de llevarlo a una forma escalonada. ¿Hay alguna ecuación inconsistente? ¿Cuántas soluciones tiene este sistema y porqué?

$$\begin{aligned} x + 6y \qquad &= \ \ 9 \\ - \ \ y - 2z &= -7 \\ 3z &= \ \ 9 \\ - \ \ y + \ \ z &= \ \ 2 \end{aligned}$$

Problema 1.5.4 Supongamos que dado un sistema de ecuaciones, y luego de realizar operaciones elementales y aplicar el método de Gauss, se obtuvo como resultado la siguiente matriz:

$$\text{Sistema } 3 \times 5\text{:} \quad \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 1 & 1 & 1 & 2 & 2 & 3 \\ 1 & 1 & 1 & 2 & 3 & 2 \end{array}\right) \longrightarrow \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right)$$

a) Escribir las ecuaciones lineales que corresponden a lo obtenido en la matriz ampliada.
b) De acuerdo a lo obtenido, ¿es compatible ese sistema?
c) ¿Cuántas soluciones tiene? Encontrar el conjunto solución, escribiéndolo cómo suma de dos componentes, una la solución del sistema homogéneo asociado y la otra, una solución “particular” del sistema no homogéneo, e indicar la cantidad de variables dependientes e independientes que posee.

Problema 1.5.5 El método de Gauss puede aplicarse de varias formas (se pueden realizar diversas operaciones elementales, se pueden elegir filas distintas, etc.). ¿Siempre se llega al mismo conjunto solución? ¿El número de variables independientes y dependientes será el mismo? Entonces, el conjunto solución en ambos casos, ¿tendrá el mismo número de parámetros libres? ¿Tendrán las mismas variables libres? (es decir, ¿es posible que en una resolución sean x e y, y en la otra resolución sean y y z, las variables libres?).

1.6. Forma escalonada reducida de Gauss-Jordan

Es una versión del método de eliminación de Gauss que si bien incrementa el número de operaciones elementales a realizar, tiene la ventaja de llegar a una forma escalonada "especial" que permite obtener directamente las expresiones finales de las variables dependientes. Posee también otras ventajas que serán discutidas en los próximos capítulos.

Si en la aplicación del método de Gauss, una vez alcanzada la matriz escalonada se continúa el proceso hasta que en cada columna correspondiente a cada variable pivote, el único elemento no nulo sea el elemento pivote y este tenga valor 1, se obtiene la denominada forma escalonada reducida de Gauss -Jordan.

Ejemplo 1.6.1 Si en el problema **1.5.4** se continúa el proceso de eliminación hasta que todos los coeficientes por arriba de cada pivote (el primer elemento no nulo de cada fila) se reduzcan a cero, se obtiene

$$\left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right) \underset{\substack{f_1 - f_3 \\ f_2 - f_3}}{\longrightarrow} \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right)$$

$$\underset{f_1 - f_2}{\longrightarrow} \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right)$$

Las variables independientes pueden ser entonces x_2 y x_3. "Pasando" estas al término derecho del sistema obtenemos

$$\begin{aligned} x_1 &= 1 - x_2 - x_3 \\ x_4 &= 2 \\ x_5 &= -1 \end{aligned}$$

Entonces, para cada par de valores $(x_2, x_3) = (\alpha, \beta)$, la solución es $(x_1, x_2, x_3, x_4, x_5) = (1 - \alpha - \beta, \alpha, \beta, 2, -1) = (1, 0, 0, 2, -1) + \alpha(-1, 1, 0, 0, 0) + \beta(-1, 0, 1, 0, 0)$. Obtenemos así infinitas soluciones.

La matriz aumentada final en el ejemplo anterior se dice que está en *forma escalonada reducida*. La ventaja de llegar a esta forma mediante el proceso de eliminación es que permite obtener en forma directa la expresión final de las variables dependientes.

Definición.
Una matriz A de coeficientes de $m \times n$ está en *forma escalonada reducida por filas* si

1. la matriz está en la "forma escalonada por filas" y
2. el primer elemento no nulo en cada fila es 1 y es el *único coeficiente no nulo* en su columna (todos los coeficientes situados arriba han sido llevados a cero).

Este proceso de eliminación más restrictivo se denomina *reducción de Gauss-Jordan.*

Ejemplo 1.6.2 Consideremos el sistema

$$\begin{aligned} 2x + y \quad\quad - \; w \quad\quad &= 4 \\ y \quad\quad + \; w + u &= 4 \\ x \quad - z + 2w \quad\quad &= 0 \end{aligned}$$

Utilizando la matriz ampliada y aplicando operaciones elementales, se obtiene:

$$\left(\begin{array}{ccccc|c} 2 & 1 & 0 & -1 & 0 & 4 \\ 0 & 1 & 0 & 1 & 1 & 4 \\ 1 & 0 & -1 & 2 & 0 & 0 \end{array}\right) \xrightarrow[(1/2)f_1]{} \left(\begin{array}{ccccc|c} 1 & 1/2 & 0 & -1/2 & 0 & 2 \\ 0 & 1 & 0 & 1 & 1 & 4 \\ 1 & 0 & -1 & 2 & 0 & 0 \end{array}\right)$$

$$\xrightarrow[f_3-(1)f_1]{} \left(\begin{array}{ccccc|c} 1 & 1/2 & 0 & -1/2 & 0 & 2 \\ 0 & 1 & 0 & 1 & 1 & 4 \\ 0 & -1/2 & -1 & 5/2 & 0 & -2 \end{array}\right)$$

$$\xrightarrow[f_3+(1/2)f_2]{} \left(\begin{array}{ccccc|c} 1 & 1/2 & 0 & -1/2 & 0 & 2 \\ 0 & 1 & 0 & 1 & 1 & 4 \\ 0 & 0 & -1 & 3 & 1/2 & 0 \end{array}\right)$$

$$\xrightarrow[(-1)f_3]{} \left(\begin{array}{ccccc|c} 1 & 1/2 & 0 & -1/2 & 0 & 2 \\ 0 & 1 & 0 & 1 & 1 & 4 \\ 0 & 0 & 1 & -3 & -1/2 & 0 \end{array}\right)$$

$$\xrightarrow[f_1-(1/2)f_2]{} \left(\begin{array}{ccccc|c} 1 & 0 & 0 & -1 & -1/2 & 0 \\ 0 & 1 & 0 & 1 & 1 & 4 \\ 0 & 0 & 1 & -3 & -1/2 & 0 \end{array}\right)$$

La última expresión es la forma de Gauss-Jordan. Despejando z de la última ecuación en términos de w y u, luego y (de la segunda) y por último x (de la primera), se obtiene el conjunto solución

$$(x, y, z, w, u) = \{(w + (1/2)u, 4 - w - u, 3w + (1/2)u, w, u) \mid w, u \in \mathbb{R}\}$$

Si usamos la forma de vectores columna, podemos expresar el conjunto solución como

$$\left\{ \begin{pmatrix} x \\ y \\ z \\ w \\ u \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 0 \\ 0 \\ 0 \end{pmatrix} + w \begin{pmatrix} 1 \\ -1 \\ 3 \\ 1 \\ 0 \end{pmatrix} + u \begin{pmatrix} 1/2 \\ -1 \\ 1/2 \\ 0 \\ 1 \end{pmatrix} \;\middle|\; w, u \in \mathbb{R} \right\}$$

Si se ponen w y u en cero se obtiene

$$\begin{pmatrix} x \\ y \\ z \\ w \\ u \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

que es una solución particular del sistema no homogéneo.

Por otra parte, el término restante,

$$\left\{ \begin{pmatrix} x \\ y \\ z \\ w \\ u \end{pmatrix} = w \begin{pmatrix} 1 \\ -1 \\ 3 \\ 1 \\ 0 \end{pmatrix} + u \begin{pmatrix} 1/2 \\ -1 \\ 1/2 \\ 0 \\ 1 \end{pmatrix} \;\Big|\; w, u \in \mathbb{R} \right\}$$

es **el conjunto solución del sistema homogéneo**.

Problemas 1.6

1. En cada caso, analizar si el sistema de ecuaciones lineales es compatible. En el caso de que sea compatible, determinar el conjunto solución y la cantidad de variables dependientes e independientes. Obtener también la forma reducida de Gauss-Jordan.

(a) $\begin{cases} 2x + 3y = 13 \\ x - y = -1 \end{cases}$ **(b)** $\begin{cases} x - z = 0 \\ 3x + y = 1 \\ -x + y + z = 4 \end{cases}$ **(c)** $\begin{cases} 2x + 2y = -4 \\ -x - y = 2 \end{cases}$

(d) $\begin{cases} -x + y = 1 \\ x + y = 2 \end{cases}$ **(e)** $\begin{cases} x - 3y + z = 1 \\ x + y + 2z = 14 \end{cases}$ **(f)** $\begin{cases} -x - y = 1 \\ -3x - 3y = 2 \end{cases}$

(g) $\begin{cases} 4y + z = 20 \\ 2x - 2y + z = 0 \\ x + z = 5 \\ x + y - z = 10 \end{cases}$ **(h)** $\begin{cases} 2x + z + w = 5 \\ y - w = -1 \\ 3x - z - w = 0 \\ 4x + y + 2z + w = 9 \end{cases}$

2. Analizar si existen valores de b y k para los cuales el sistema es i) incompatible ii) compatible determinado y iii) compatible indeterminado, y hallar el conjunto solución en los casos compatibles.

(a) $\begin{cases} x - y = 1 \\ 3x - 3y = b \end{cases}$ **(b)** $\begin{cases} x - y = 1 \\ 3x + ky = 3 \end{cases}$ **(c)** $\begin{cases} x - y = 1 \\ 3x + ky = b \end{cases}$

3. a) ¿Qué condiciones deben cumplir las constantes b_i para que cada sistema tenga solución? i) $\begin{cases} x - 3y = b_1 \\ 3x + \ \ y = b_2 \\ x + 7y = b_3 \\ 2x + 4y = b_4 \end{cases}$ ii) $\begin{cases} x_1 + 2x_2 + 3x_3 = b_1 \\ 2x_1 + 5x_2 + 3x_3 = b_2 \\ x_1 \qquad\quad + 8x_3 = b_3 \end{cases}$

   b) Sin realizar cálculos adicionales, analizar los sistemas homogéneos asociados (es decir, cuando $b_i = 0 \ \forall \ i$).

4. Analizar el sistema según los valores de k y h. En caso de ser compatible (determinado o indeterminado) dar el conjunto solución.

$$\begin{cases} x + y + \ \ z = h \\ x + y + kz = 1 \end{cases}$$

5. Encuentre, de ser posible, coeficientes a, b, y c tales que el gráfico de $f(x) = ax^2 + bx + c$ pase por los puntos $(1, 2)$, $(-1, 6)$, y $(2, 3)$. ¿Son únicos?

6. Mostrar que si $ad - bc \neq 0$, el sistema

$$\begin{aligned} ax + by &= b_1 \\ cx + dy &= b_2 \end{aligned}$$

   posee solución **única** $\forall \ b_1, b_2$, mientras que si $ad - bc = 0$, el sistema es o bien incompatible o bien compatible indeterminado.

7. Resolver los sistemas presentados en la introducción del capítulo: problema de la palanca, flujo de redes, problema de temperaturas y problema altimétrico, verificando los resultados indicados. Discutir las soluciones encontradas en cada caso interpretando los resultados obtenidos en el contexto del problema. En el caso del problema altimétrico, discutir posibles soluciones, según sean los valores de los desniveles $\Delta H_{12}, \Delta H_{23}, \Delta H_{31}$. En el caso de flujo de redes, considerar que un flujo negativo en alguna rama representa un cambio de dirección del mismo.

8. Resolver los sistemas del problema 1.6.2 mediante un software adecuado (por ej., mathlab, mathematica, maple, etc.). Nótese que los programas para resolver sistemas lineales utilizan justamente métodos matriciales basados en la reducción por filas.

9. Tres conductores se detuvieron en un bar del camino.
   Uno de ellos compró cuatro sándwiches, una taza de café y diez medialunas, pagando un total de \$150. Otro conductor compró tres sándwiches, una taza de café y siete medialunas, pagando \$120.
   ¿Es posible saber cuanto pagó un tercer conductor por un sándwich, una taza de café y una medialuna?
   ¿Es posible conocer el costo unitario de cada producto?

10. Se tiene un conjunto de n números $(x_1, \dots, x_n)$ tales que los números interiores son el promedio de los números vecinos ($x_i = \frac{x_{i+1}+x_{i-1}}{2}$ para $i = 2, \dots, n-1$), y $x_n - x_1 = n - 1$, con $x_1 = t$. Dé una expresión para los x_i en función de t y n. (Sugerencia: considere primero $n = 3$).

11. **Para pensar**
    Varios *sistemas de ecuaciones lineales* que surgen de *modelar* problemas que involucran datos experimentales, como por ejemplo el problema altimétrico, con frecuencia son *incompatibles* o *inconsistentes* y a veces también "mal condicionados". Es decir, no tienen solución o si la tienen la misma es muy "sensible" a cambios o perturbaciones en los datos del problema. En el caso de sistemas inconsistentes se recurre a un método conocido por *Método de Mínimos Cuadrados* que permite encontrar la «mejor» solución al sistema de ecuaciones lineales. Estos temas se estudiarán en la Parte II.

Capítulo 2

Matrices

2.1. Introducción

En el capítulo previo hemos introducido el concepto de matriz para representar los coeficientes de los sistemas de ecuaciones lineales. En este estudiaremos operaciones entre matrices y las aplicaremos a la resolución de dichos sistemas. En particular, veremos que el análisis y la resolución de sistemas podrá ser realizado por medio de las propiedades algebraicas de las matrices que los representan. El *algebra matricial* permitirá también expresar los sistemas en forma en forma concisa, lo que resulta muy adecuado para el estudio de propiedades generales. La matrices son además utilizadas para manipular y correlacionar datos, y para representar operaciones geométricas y matemáticas (Cap. V).

La introducción y uso de matrices se remonta a siglos, incluso milenios, atrás, siendo que de épocas ciertamente remotas (siglo VII a.C.) datan los estudios de los llamados *cuadrados mágicos.* El uso de matrices, y determinantes (concepto que será introducido en el siguiente capítulo) para sistemas de ecuaciones lineales fue presentado en el siglo XVII por Gottfried Leibnitz y Seki Kowa, mientras que importantes contribuciones posteriores fueron aportadas por Carl F. Gauss y Wilhelm Jordan, quienes desarrollaron el método visto en el capítulo de sistemas lineales. Matemáticos como William Hamilton, Arthur Cayley, James Sylvester, John von-Neumann, Camille Jordan, entre otros, trabajaron en temas de Algebra Matricial. Cabe destacar además la formulación matricial realizada originalmente por Werner Heisenberg de la rama de Física llamada Mecánica Cuántica.

2.2. Conceptos básicos

Definición.
Una *matriz* A de $m \times n$ es un arreglo rectangular de $m \times n$ números a_{ij}, $i = 1, ..., m$, $j = 1, ..., n$, dispuestos en m filas y n columnas:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \tag{2.1}$$

De esta manera, el primer subíndice, i, indica la fila y el segundo, j, la columna en la que se encuentra el elemento a_{ij}. $m \times n$ es la *dimensión* de la matriz.

Otras notaciones utilizadas para denotar una matriz A de elementos a_{ij} es (a_{ij}) o $[a_{ij}]$. También se emplea directamente la notación A_{ij} o $A_{i,j}$ para denotar el elemento de la fila i y columna j de una matriz A ($A_{ij} = a_{ij}$ en (2.1)). Por ejemplo,

$$A = \begin{pmatrix} 1 & -2 & 5 \\ 9 & 0 & 3 \end{pmatrix}$$

es una matriz de 2×3. En este caso, $a_{12} = -2$, $a_{21} = 9$.

Cuando $m = n$ se dice que la matriz es *cuadrada.* Por ejemplo,

$$B = \begin{pmatrix} 1 & -2 \\ 9 & 0 \end{pmatrix}$$

es una matriz de 2×2. B es una *submatriz* de A, formada por sus dos primeras columnas.

El conjunto de todas las matrices de $m \times n$ con elementos reales se denota $\mathbb{R}^{m\times n}$. Si los elementos son números complejos, el conjunto se denota $\mathbb{C}^{m\times n}$.

Si $m = n = 1$, $\mathbb{R}^{1\times 1}$ se considera equivalente al conjunto $\mathbb{R}$ de números reales. Por otro lado, cuando $m = 1$ o $n = 1$, la matriz resulta equivalente a un vector de n o m componentes respectivamente:

Definición.
Un vector o matriz columna $\boldsymbol{b}$ de m componentes es una matriz de dimensión $m \times 1$:

$$\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

Un vector o matriz fila $\mathbf{a}$ de n componentes es una matriz de dimensión $1 \times n$:

$$\mathbf{a} = (a_1, a_2, \ldots, a_n)$$

2.2.1. Operaciones básicas con matrices

Las siguientes son definiciones básicas del álgebra de matrices.

Dos matrices A y B se dice que son *iguales* si y sólo si tienen:
i) *la misma dimensión* (ambas de $m \times n$) y
ii) *todos sus elementos correspondientes iguales*: $a_{ij} = b_{ij}\ \forall\ i, j$.

Multiplicación por un escalar.
Sea A una matriz de $m \times n$ y α un escalar (un número real o complejo). El producto αA es una matriz de igual dimensión cuyos elementos son los de A multiplicados por α:

$$(\alpha A)_{ij} = \alpha a_{ij}$$

Por definición $A\alpha = \alpha A$ (α escalar).

Es decir, *todos* los elementos de A deben ser multiplicados por α. Por ejemplo,

$$3\begin{pmatrix} 1 & -2 & 5 \\ 9 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 3 & -6 & 15 \\ 27 & 0 & 9 \end{pmatrix}$$

Suma de matrices.
Sean A y B dos matrices *de la misma dimensión* (ambas de $m \times n$). La suma $A + B$ es la matriz de igual dimensión cuyos elementos son la suma de los correspondientes elementos de A y B:

$$(A + B)_{ij} = a_{ij} + b_{ij}$$

Si las dimensiones no son iguales la suma **no está definida**.

Por ejemplo,

$$\begin{pmatrix} 1 & -2 \\ 9 & 0 \end{pmatrix} + \begin{pmatrix} 3 & 4 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 8 & 2 \end{pmatrix}$$

Propiedades de la suma de matrices.
Las siguientes propiedades son obvias a partir de la definición.

1. **Conmutatividad**. Para todo par de matrices A, B de $m \times n$, se cumple

$$A + B = B + A$$

2. **Asociatividad**. Para toda terna de matrices A, B, C de $m \times n$, se cumple

$$A + (B + C) = (A + B) + C$$

3. **Existencia de elemento neutro (matriz nula)**. Para toda A de $m \times n$ se cumple

$$A + \mathbf{0} = \mathbf{0} + A = A$$

donde

$$\mathbf{0} = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}$$

es la **matriz nula** de $m \times n$ ($\mathbf{0}_{ij} = 0 \ \forall \ i, j$).

4. **Existencia de elemento opuesto** $(-A)$**.** Para toda A de $m \times n$ se cumple

$$A + (-A) = (-A) + A = \mathbf{0}$$

donde

$$-A = (-1)A = \begin{pmatrix} -a_{11} & \dots & -a_{1n} \\ \vdots & \ddots & \vdots \\ -a_{m1} & \dots & -a_{mn} \end{pmatrix}$$

La **resta** de dos matrices A, B de $m \times n$ se define entonces como

$$A - B = A + (-B)$$

Por ejemplo, $\begin{pmatrix} 1 & -2 \\ 9 & 0 \end{pmatrix} - \begin{pmatrix} 1 & -2 \\ 9 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

En relación a la multiplicación por un escalar, es fácil probar las propiedades siguientes:

5. **Distributividad respecto de la suma de matrices**. Si A y B tienen la misma dimensión, para todo escalar α se cumple

$$\alpha(A + B) = \alpha A + \alpha B$$

6. **Distributividad respecto de la suma de escalares**. Si α y β son escalares, para toda matriz A se cumple

$$(\alpha + \beta)A = \alpha A + \beta A$$

7. **Asociatividad.** Para toda matriz A y escalares α, β, se cumple

$$(\alpha\beta)A = \alpha(\beta A)$$

Al intercambiar (o trasponer) filas por columnas en una matriz A se obtiene la denominada *matriz traspuesta* (o transpuesta) A^T:

Matriz traspuesta.
La traspuesta A^T de una matriz A de $m \times n$ es una matriz de $n \times m$ cuyas filas son las columnas de A, es decir, cuyo elemento i, j es el elemento j, i de A:

$$(A^T)_{ij} = a_{ji} \tag{2.2}$$

Por ejemplo,

$$\begin{pmatrix} 1 & -2 & 5 \\ 9 & 0 & 3 \end{pmatrix}^T = \begin{pmatrix} 1 & 9 \\ -2 & 0 \\ 5 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 9 \\ -2 & 0 \\ 5 & 3 \end{pmatrix}^T = \begin{pmatrix} 1 & -2 & 5 \\ 9 & 0 & 3 \end{pmatrix}$$

Obviamente, por la definición se cumple

$$(A^T)^T = A \tag{2.3}$$

Si A y B son de $m \times n$ y α es un escalar, se cumplen también las siguientes propiedades:

$$(A + B)^T = A^T + B^T \tag{2.4}$$

$$(\alpha A)^T = \alpha A^T \tag{2.5}$$

es decir, la traspuesta de una suma de matrices es la suma de sus traspuestas, y la traspuesta de un multiplo de una matriz es el multiplo de su traspuesta.
La demostración de estas propiedades (obvias) se dejan para el lector. Por ejemplo, $((A + B)^T)_{ij} = (A + B)_{ji} = a_{ji} + b_{ji} = (A^T)_{ij} + (B^T)_{ij} = (A^T + B^T)_{ij}$, prueba (2.4).

Nótese que la traspuesta de una matriz fila es una matriz columna y viceversa:

$$\begin{pmatrix} 1 \\ -2 \\ 5 \end{pmatrix}^T = \begin{pmatrix} 1 & -2 & 5 \end{pmatrix}, \quad \begin{pmatrix} 1 & -2 & 5 \end{pmatrix}^T = \begin{pmatrix} 1 \\ -2 \\ 5 \end{pmatrix}$$

2.2.2. Matrices cuadradas especiales

Como se mencionó previamente, si $m = n$ la matriz se dice cuadrada. Los siguientes tipos de matrices cuadradas son de particular importancia, como veremos luego.

Matriz diagonal
Es aquella en la que todos los elementos que están fuera de la diagonal principal son cero:

$$A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & a_{nn} \end{pmatrix}$$

es decir, $a_{ij} = 0$ si $i \neq j$.

Matriz identidad
Es una matriz diagonal donde todos los elementos de la diagonal principal son 1. Se la denota con I o también $\mathbb{1}$:

$$I \;=\; \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.6}$$

es decir, $I_{ij} = \left\{ \begin{array}{ll} 1 & i = j \\ 0 & i \neq j \end{array} \right.$. La matriz identidad de $n \times n$ se denota como I_n o $\mathbb{1}_n$.

Matriz triangular superior
Es una matriz donde todos los coeficientes por debajo de la diagonal principal son cero:

$$A \;=\; \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & 0 & a_{nn} \end{pmatrix}$$

es decir, $a_{ij} = 0$ si $i > j$.
Matriz triangular inferior.
Es una matriz donde todos los coeficientes por encima de la diagonal principal son cero:

$$A \;=\; \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

es decir, $a_{ij} = 0$ si $i < j$. Si A es triangular superior, A^T es triangular inferior y viceversa.

Matriz simétrica. Es una matriz que es igual a su matriz traspuesta:

$$A \text{ simétrica } \Leftrightarrow \; A^T = A$$

o sea $a_{ij} = a_{ji} \; \forall \; i, j$. Por ejemplo,

$$A = \begin{pmatrix} 1 & -7 & 3 \\ -7 & 2 & 0 \\ 3 & 0 & -4 \end{pmatrix}$$

Matriz anti-simétrica. Es una matriz que es igual a menos su matriz traspuesta:

$$A \text{ antisimétrica } \Leftrightarrow \; A^T = -A$$

o sea $a_{ij} = -a_{ji} \; \forall \; i, j$. Esto implica que los elementos diagonales $(i = j)$ son nulos, ya que $A_{ii} = -A_{ii}$. Por ejemplo,

$$A = \begin{pmatrix} 0 & -7 & 3 \\ 7 & 0 & 1 \\ -3 & 1 & 0 \end{pmatrix}$$

Problemas 2.2

1. Dada $A=\begin{pmatrix}1 & 3 & 1\\ 2 & -1 & 4\end{pmatrix}$, determinar:

   a) Los elementos a_{12} y a_{21} , b) $2A$, c) $0A$, d) A^T
   e) ¿Está definida la suma $A+A^T$? ¿Como debe ser la matriz A para que la suma $A+A^T$ esté definida?

2. Realizar las siguientes operaciones, en caso de estar definidas.

   (a) $3\begin{pmatrix}2\\1\\1\end{pmatrix}+2\begin{pmatrix}3\\0\\4\end{pmatrix}$ **(b)** $5\begin{pmatrix}4\\-1\end{pmatrix}+2\begin{pmatrix}1\\2\\3\end{pmatrix}$ **(c)** $2\begin{pmatrix}1 & 3\\ 2 & -1\end{pmatrix}+3\begin{pmatrix}3 & 2\\ 1 & 1\end{pmatrix}$

   (d) $2\begin{pmatrix}1 & 3 & 1\end{pmatrix}^T+3\begin{pmatrix}1\\2\\3\end{pmatrix}$ **(e)** $2\begin{pmatrix}1 & 3\\ 2 & 4\end{pmatrix}^T-\begin{pmatrix}1 & 2\\ 2 & -1\end{pmatrix}$

3. Determinar, en caso de que existan, todos los α y β tales que

   (a) $\beta\begin{pmatrix}5\\-5\end{pmatrix}+\alpha\begin{pmatrix}1\\-1\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$ **(b)** $\begin{pmatrix}-1\\2\\1\end{pmatrix}+\alpha\begin{pmatrix}-2\\1\\0\end{pmatrix}+\beta\begin{pmatrix}3\\0\\1\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$

4. Determinar, en caso de que existan, todas las matrices B que satisfacen

   (a) $\begin{pmatrix}1 & 2\\ 3 & -1\end{pmatrix}+B=B^T+\begin{pmatrix}1 & 1\\ 4 & -1\end{pmatrix}$ **(b)** $\begin{pmatrix}1 & 2\\ 3 & -1\end{pmatrix}+B=2B^T+\begin{pmatrix}1 & 1\\ 4 & -1\end{pmatrix}$

5. Demostrar que si A es de $n\times n$, entonces $A+A^T$ es siempre una matriz **simétrica**, y $A-A^T$ es siempre una matriz **antisimétrica**. Dar un ejemplo de 2×2.

6. Mostrar que toda matriz A de $n\times n$ puede escribirse como $A=A_s+A_a$, con A_s simétrica y A_a antisimétrica. Mostrar también que A_s y A_a son únicas.

7. En la matriz siguiente se disponen las calificaciones de 5 estudiantes, obtenidas en 3 exámenes (puntaje máximo = 10 en cada uno). Cada columna corresponde al resultado de cada examen, mientras que las filas corresponden a los estudiantes.

   $$\text{Estudiantes}\;\overset{\text{Exámenes}}{\begin{pmatrix}7 & 6 & 8.5\\ 9 & 9.5 & 10\\ 6 & 7 & 6.5\\ 6 & 8 & 4\\ 7.5 & 7 & 7\end{pmatrix}}=A$$

   (i) Si las calificaciones son modificadas agregando a todos los alumnos 1 punto a las del primer examen y .5 puntos a las del segundo examen, encontrar, usando la suma de matrices, una forma de calcular las nuevas calificaciones.
   (ii) Si se decide reducir un 10 % todas las notas, encuentre una forma de realizar esta operación en forma matricial (sugerencia: multiplique por un escalar adecuado).
   (iii) El profesor desea computar los promedios finales, considerando que el promedio proviene de la siguiente ponderación: 30 % del primer examen, 30 % del segundo y 40 % del tercero. Pensarlo como suma de tres vectores columnas.

   (iv) Una vez determinada la forma matricial, realizar los cálculos mediante PC o equivalente utilizando algún software adecuado.

2.3. Producto de matrices

Pasemos ahora a definir el producto de matrices. Esta es una operación clave, que posibilita el uso de las matrices para representar algebraicamente sistemas de ecuaciones lineales, y también, como se verá en el Cap. V, transformaciones lineales entre vectores y operaciones geométricas tales como rotaciones y reflexiones.

Antes de definir el producto matricial, recordemos que **el producto escalar** o punto de dos vectores reales de n componentes $\boldsymbol{a} = (a_1, \ldots, a_n)$ y $\boldsymbol{b} = (b_1, \ldots, b_n)$ está dado por

$$\boldsymbol{a} \cdot \boldsymbol{b} = \sum_{k=1}^{n} a_k b_k = a_1 b_1 + \ldots + a_n b_n$$

Definición

Consideremos una matriz A de $m \times n$ y una matriz B de $n \times p$, tal que el número de columnas de A coincide con el número de filas de B. Las filas de A y las columnas de B,

$$\boldsymbol{a}_{i*} = (a_{i1}, \ldots, a_{in}), \quad \boldsymbol{b}_{*j} = \begin{pmatrix} b_{1j} \\ \ldots \\ b_{nj} \end{pmatrix}$$

son entonces vectores de n componentes. El producto de A por B es una matriz de $m \times p$ cuyos elementos i, j son el **producto escalar** de la fila i de A por la columna j de B:

A	B	AB
$m \times n$	$n \times p$	$m \times p$

$$\begin{aligned} (AB)_{ij} &= \boldsymbol{a}_{i*} \cdot \boldsymbol{b}_{*j} \\ &= \sum_{k=1}^{n} a_{ik} b_{kj} = a_{i1} b_{1j} + \ldots + a_{in} b_{nj} \end{aligned}$$

Es decir,

$$AB = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & \ldots & b_{1p} \\ \vdots & \ddots & \vdots \\ b_{n1} & \ldots & b_{np} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{n} a_{1k} b_{k1} & \ldots & \sum_{k=1}^{n} a_{1k} b_{kp} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{n} a_{mk} b_{k1} & \ldots & \sum_{k=1}^{n} a_{mk} b_{kp} \end{pmatrix}$$

Si el número de columnas de A no coincide con el número de filas de B, el producto AB **no está definido**.

Ejemplos 2.3.1:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1+4 & -2+2 \\ 3+8 & -6+4 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 11 & -2 \end{pmatrix} \tag{2.7}$$

$$\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1-6 & 2-8 \\ 2+3 & 4+4 \end{pmatrix} = \begin{pmatrix} -5 & -6 \\ 5 & 8 \end{pmatrix} \tag{2.8}$$

¡¡Se observa que el orden de los factores **sí altera el producto!!**

Además, el producto de una matriz fila de $1 \times n$ por una matriz columna de $n \times 1$ es una matriz de 1×1, mientras que el de una matriz columna de $n \times 1$ por una matriz fila de $1 \times n$ es una matriz de $n \times n$!! Por ejemplo, para $n = 2$,

$$\begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 7 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 3 \end{pmatrix} \begin{pmatrix} 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} \tag{2.9}$$

En el caso general, AB puede estar definido pero BA no necesariamente lo estará, como muestra el siguiente ejemplo con A de 2×3 y B de 3×1:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 4 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 4 \end{pmatrix} \text{ no definido} \tag{2.10}$$

Importante: No conmutatividad del producto matricial
¡¡Como muestran los ejemplos anteriores, el producto matricial **no es conmutativo!!**

- Si A y B son matrices cuadradas de $n \times n$, AB y BA están ambos definidos y tienen la misma dimensión ($n \times n$), pero salvo casos especiales, en general

$$AB \neq BA \tag{2.11}$$

  como sucede en (2.7)–(2.8).
- Si A es de $m \times n$ y B de $n \times m$, con $m \neq n$, AB y BA siguen estando definidos pero ya no tienen la misma dimensión: AB será de $m \times m$ y BA de $n \times n$, tal como muestra (2.9) para $n = 2$, $m = 1$. En este caso $AB \neq BA$ siempre.
- Si A es de $m \times n$ y B de $n \times p$, con $p \neq m$, AB estará definido pero BA **no estará definido**, tal como muestra (2.10) para $m = 2$, $n = 3$, $p = 1$.

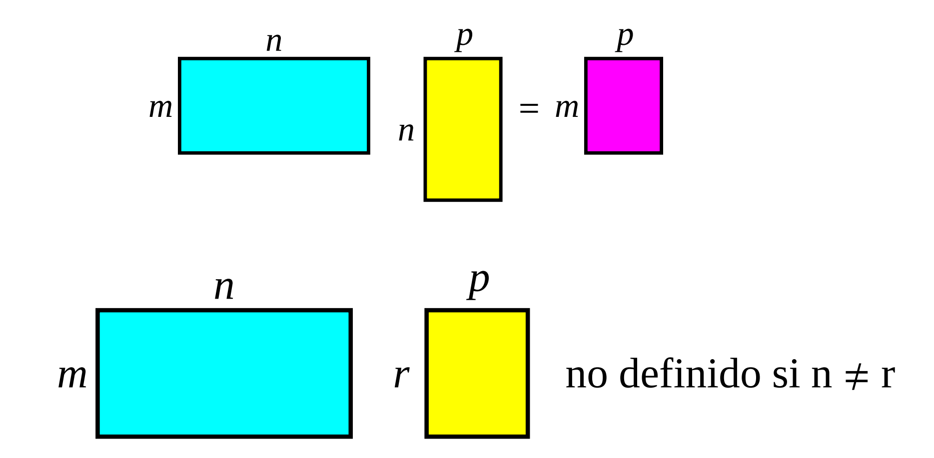


Figura 2.1: Esquema del producto matricial. Arriba el caso definido y abajo el no definido.

No obstante, son válidas las siguientes propiedades:

- **Asociatividad.** Si A es de $m \times n$, B de $n \times p$ y C de $p \times q$, entonces

$$A(BC) = (AB)C \tag{2.12}$$

por lo que el producto anterior se escribe simplemente ABC (de $m \times q$).
Demostración:

$$\begin{aligned}(A(BC))_{ij} &= \sum_{k=1}^{n} a_{ik}(BC)_{kj} = \sum_{k=1}^{n} a_{ik}(\sum_{l=1}^{p} b_{kl}c_{lj}) = \sum_{l=1}^{p}(\sum_{k=1}^{n} a_{ik}b_{kl})c_{lj} = \sum_{l=1}^{p}(AB)_{il}c_{lj} \\ &= ((AB)C)_{ij}\end{aligned}$$

Ejemplo:

$$(1\ 2)\left(\begin{pmatrix}1 & 2\\3 & 4\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}\right) = (1\ 2)\begin{pmatrix}3\\7\end{pmatrix} = (17),\ \left((1\ 2)\begin{pmatrix}1 & 2\\3 & 4\end{pmatrix}\right)\begin{pmatrix}1\\1\end{pmatrix} = (7\ 10)\begin{pmatrix}1\\1\end{pmatrix} = (17)$$

- **Distributividad.**
Si A es de $m \times n$ y B, C de $n \times p$,

$$A(B + C) = AB + AC$$

Si A, B son de $m \times n$ y C de $n \times p$,

$$(A + B)C = AC + BC$$

La demostración de estas dos propiedades se deja para el lector. Por ej.,
$(A(B + C))_{ij} = \boldsymbol{a}_{i*} \cdot (\boldsymbol{b}_{*j} + \boldsymbol{c}_{*j}) = \boldsymbol{a}_{i*} \cdot \boldsymbol{b}_{*j} + \boldsymbol{a}_{i*} \cdot \boldsymbol{c}_{*j} = (AB)_{ij} + (AC)_{ij}$.

- **Asociatividad con el producto por escalar**. Si A es de $m \times n$ y B de $n \times p$,

$$\alpha(AB) = (\alpha A)B = A(\alpha B)$$

para todo escalar α. La demostración (muy fácil) se deja para el lector.

Traspuesta de un producto
Si A es de $m \times n$ y B de $n \times p$,

$$(AB)^T = B^T A^T \qquad (p \times m)$$

Notar que **se invierte el orden**. B^T es de $p \times n$ y A^T de $n \times m$, por lo que $B^T A^T$ es de $p \times m$ como $(AB)^T$ y está siempre definido si AB lo está (a diferencia de $A^T B^T$).
Demostración:

$$((AB)^T)_{ij} = (AB)_{ji} = \sum_{k=1}^{n} a_{jk}b_{ki} = \sum_{k=1}^{n} b_{ki}a_{jk} = \sum_{k=1}^{n}(B^T)_{ik}(A^T)_{kj} = (B^T A^T)_{ij}$$

En forma concisa, si $\boldsymbol{b}^T_{i*}$ denota la fila i de B^T y $\boldsymbol{a}^T_{*j}$ la columna j de A^T,

$$(AB)^T_{ij} = \boldsymbol{a}_{j*} \cdot \boldsymbol{b}_{*i} = \boldsymbol{b}^T_{i*} \cdot \boldsymbol{a}^T_{*j} = (B^T A^T)_{ij}$$

Ejemplo 2.3.2

$$\left(\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}\begin{pmatrix}1 & -2\\ 2 & 1\end{pmatrix}\right)^T = \begin{pmatrix}5 & 0\\ 11 & -2\end{pmatrix}^T = \begin{pmatrix}5 & 11\\ 0 & -2\end{pmatrix} = \begin{pmatrix}1 & 2\\ -2 & 1\end{pmatrix}\begin{pmatrix}1 & 3\\ 2 & 4\end{pmatrix}$$

Observación 1. Producto escalar como producto matricial
Si A es una matriz fila $\boldsymbol{a}$ de $1 \times n$ y B una matriz columna $\boldsymbol{b}$ de $n \times 1$, AB es una matriz de 1×1 cuyo único elemento es el producto escalar $\boldsymbol{a} \cdot \boldsymbol{b}$:

$$AB = \begin{pmatrix}a_1 & \dots & a_n\end{pmatrix}\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix} = a_1 b_1 + \ldots + a_n b_n$$

(si la matriz es de 1×1 no es necesario escribir los paréntesis: se identifica $\mathbb{R}^{1\times 1}$ con $\mathbb{R}$).
Ejemplo:

$$\begin{pmatrix}1 & 2 & 3\end{pmatrix}\begin{pmatrix}2\\ -1\\ 1\end{pmatrix} = 3$$

Observación 2. Otras peculiaridades del producto matricial

- $AB = 0$ (matriz nula) no implica $A = 0$ o $B = 0$, ni tampoco $BA = 0$. Por ejemplo,

$$\begin{pmatrix}1 & 1\\ 2 & 2\end{pmatrix}\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix} = \begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix} \text{ pero } \begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}\begin{pmatrix}1 & 1\\ 2 & 2\end{pmatrix} = \begin{pmatrix}-1 & -1\\ 1 & 1\end{pmatrix}$$

- $AB = AC$ no implica $B = C$, aún si $A \neq 0$. Por ejemplo,

$$\begin{pmatrix}1 & 1\\ 2 & 2\end{pmatrix}\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix} = \begin{pmatrix}4 & 6\\ 8 & 12\end{pmatrix} = \begin{pmatrix}1 & 1\\ 2 & 2\end{pmatrix}\begin{pmatrix}2 & 1\\ 2 & 5\end{pmatrix}$$

Observación 3. Potencias de matrices cuadradas
Si A es de $n \times n$, se define la **potencia** A^k (de $n \times n$) para todo natural $k = 1, 2, \ldots$:

$$A^2 = AA, \quad A^3 = AA^2 = A^2 A = AAA, \quad \ldots \quad A^k = AA^{k-1} = A^{k-1}A = \underbrace{AA\ldots A}_{k \text{ veces}}$$

Si, en cambio, A no es cuadrada, A^2 (y por ende cualquier potencia) **no está definida**.

Ejemplo 2.3.3 Si $A = \begin{pmatrix}1 & 2\\ 2 & 0\end{pmatrix}$,

$$A^2 = \begin{pmatrix}1 & 2\\ 2 & 0\end{pmatrix}\begin{pmatrix}1 & 2\\ 2 & 0\end{pmatrix} = \begin{pmatrix}5 & 2\\ 2 & 4\end{pmatrix}, \quad A^3 = AA^2 = \begin{pmatrix}9 & 10\\ 10 & 4\end{pmatrix}$$

Nótese que en general, $(A^k)_{ij} \neq a_{ij}^k$ si A no es diagonal.

En cambio, si A es diagonal, puede fácilmente mostrar el lector que

$$A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & a_{nn} \end{pmatrix} \Rightarrow A^k = \begin{pmatrix} a_{11}^k & 0 & \dots & 0 \\ 0 & a_{22}^k & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & a_{nn}^k \end{pmatrix}$$

Por ejemplo, $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}^2 = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix} = \begin{pmatrix} 2^2 & 0 \\ 0 & 3^2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}^k = \begin{pmatrix} 2^k & 0 \\ 0 & 3^k \end{pmatrix}$

Observación 4. Producto de matriz por vector columna
Si A es de $m \times n$ y $\mathbf{x}$ de $n \times 1$, $A\mathbf{x}$ es un vector columna de $m \times 1$, que puede expresarse como suma de las columnas de A multiplicadas por los elementos de $\mathbf{x}$:

$$A\mathbf{x} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix} \quad (2.13)$$

$$= x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \quad (2.14)$$

La última expresión es una **combinación lineal** de las columnas de A. Estos resultados serán utilizados para representar sistemas de ecuaciones lineales. Por ejemplo,

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & -4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 2y + 3z \\ 2x + 0y - 4z \end{pmatrix} = x \begin{pmatrix} 1 \\ 2 \end{pmatrix} + y \begin{pmatrix} 2 \\ 0 \end{pmatrix} + z \begin{pmatrix} 3 \\ -4 \end{pmatrix}$$

Problemas 2.3

1. En caso de estar definidas, evaluar las siguientes operaciones:

a) $\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$ b) $\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$ c) $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 1 \end{pmatrix}$ d) $\begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix}$

e) $\begin{pmatrix} 1 & 2 & 3 \end{pmatrix}\begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$ f) $\begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$ g) $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ h) $\begin{pmatrix} x \\ y \\ z \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix}$

i) $\begin{pmatrix} -1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 & 0 \\ -1 & 1 & 0 \end{pmatrix}$ j) $\begin{pmatrix} 1 & 2 & 0 \\ -1 & 1 & 0 \end{pmatrix}^T \begin{pmatrix} -1 & 2 \end{pmatrix}^T$ k) $\begin{pmatrix} 1 & 1 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

2. Muestre que si A es de $n \times n$,
   a) $(A^2)^T = (A^T)^2$, y en general, $(A^k)^T = (A^T)^k$.
   b) $A^{k+m} = A^k A^m$; c) $(\alpha A)^k = \alpha^k A^k$

3. Si A y B son de $n \times n$, exprese $(AB^2)^T$ en términos de A^T y B^T.

4. Demuestre que $\forall$ matriz A se cumple que $A^T A$ y AA^T están siempre definidas y son matrices **simétricas**. Indique sus dimensiones si A es de $m \times n$. Verifique el resultado para una matriz A de 2×2 no simétrica.

5. Si A y B son de $n \times n$, dé una expresión para $(A+B)(A-B)$ y muestre que no es necesariamente igual a $A^2 - B^2$. Dé un ejemplo.

6. Muestre que $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ 3 \end{pmatrix} + y \begin{pmatrix} 2 \\ 4 \end{pmatrix}$.

7. Exprese los promedios en el problema 2.2.7 como el producto de una matriz por un vector columna adecuado.

8. Tres personas (A, B, y C) trabajan para una empresa que produce 3 tipos de productos: P_1, P_2, P_3. La labor se paga por cada unidad realizada, dependiendo ese valor del tipo de producto. Los valores pagados son $x_1 = 100\$$ por cada unidad de P_1, $x_2 = 200\$$ por las unidades de P_2, y $x_3 = 300\$$ por cada unidad de P_3.

   Las matrices L y M siguientes representan las unidades producidas de cada producto por cada persona, durante dos días (lunes y martes por ejemplo).

   $L =$

	P_1	P_2	P_3
A	4	3	2
B	5	1	2
C	3	4	1

   $M =$

	P_1	P_2	P_3
A	3	6	1
B	4	2	2
C	5	1	3

   El vector columna o matriz X de 3×1 es el pago por cada unidad producida:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 100 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

   Calcular las matrices siguientes, y explicar su significado:
   (a) LX, (b) MX, (c) $L+M$, (d) $(L+M)X$.
   (e) Realizar las operaciones anteriores utilizando un software adecuado.

2.4. Representación matricial de sistemas lineales

Consideremos nuevamente un sistema de m ecuaciones lineales con n incógnitas $x_1, \ldots, x_n$:

$$\left\{\begin{array}{ccc} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = & b_2 \\ \vdots & \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & b_m \end{array}\right. \tag{2.15}$$

Existen dos formas equivalentes de representar este sistema en forma matricial.

En términos de la matriz de coeficientes A del sistema, de $m \times n$,

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \tag{2.16}$$

y los vectores columna $\mathbf{x}$ de incógnitas $(n \times 1)$ y $\boldsymbol{b}$ de términos independientes $(m \times 1)$,

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \tag{2.17}$$

podemos escribir el sistema (2.15) en forma matricial como (ver (2.13) en observación 4)

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \tag{2.18}$$

o sea, en forma compacta,

$$A\mathbf{x} = \boldsymbol{b} \tag{2.19}$$

En efecto, $A\mathbf{x}$ es un vector columna de $m \times 1$ cuya fila i es el miembro izquierdo de la ecuación i de (2.15) (que es igual al producto escalar $\boldsymbol{a}_{i*} \cdot \mathbf{x}$), el cual debe ser igual a b_i, es decir, a la fila i de $\boldsymbol{b}$.

Ejemplo 2.4.1 El sistema

$$\left\{\begin{array}{ccc} x_1 + 3x_2 + x_3 & = & 5 \\ x_1 + 2x_2 + 3x_3 & = & -2 \\ x_1 + 2x_2 + 2x_3 & = & 1 \end{array}\right. \tag{2.20}$$

puede escribirse como

$$\begin{pmatrix} 1 & 3 & 1 \\ 1 & 2 & 3 \\ 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ -2 \\ 1 \end{pmatrix} \tag{2.21}$$

La segunda forma de expresar el sistema (2.15) es utilizando la expresión (2.14):

Podemos escribir $A\mathbf{x}$ como combinación lineal de las columnas de A y por lo tanto, (2.18) es equivalente a

$$x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \ldots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \tag{2.22}$$

es decir,

$$x_1 \boldsymbol{a}_{*1} + \ldots + x_n \boldsymbol{a}_{*n} = \boldsymbol{b} \tag{2.23}$$

donde $\boldsymbol{a}_{*i}$ es la columna i de A.

Por ejemplo, el sistema (2.20) puede escribirse también como

$$x_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} 3 \\ 2 \\ 2 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ -2 \\ 1 \end{pmatrix} \tag{2.24}$$

Esta forma permite derivar en forma **inmediata** el siguiente Teorema general:

Teorema 2.4.1
Un sistema lineal $A\mathbf{x} = \boldsymbol{b}$ de $m \times n$ es compatible (es decir, tiene al menos una solución) **si y sólo si** el vector columna $\boldsymbol{b}$ puede ser expresado como combinación lineal de las columnas de A.

En efecto, si existe solución $\mathbf{x}$ de (2.15), entonces se cumple (2.18) y por lo tanto (2.22), por lo que $\boldsymbol{b}$ es combinación lineal de las columnas de A. Y si $\boldsymbol{b}$ es combinación lineal de las columnas A, es decir, si existen números $x_1, \ldots, x_n$ tales que se cumpla (2.22), también se cumple (2.18), por lo que $\mathbf{x}$ es solución del sistema (2.15).

Problema 2.4.1
Verificar que el sistema (2.20) tiene la solución única $\mathbf{x} = \begin{pmatrix} 5 \\ 1 \\ -3 \end{pmatrix}$ y que se cumple (2.24).

Ejemplo 2.4.2
El sistema $\left\{ \begin{array}{rcl} x + 2y & = & 1 \\ 2x + 4y & = & 1 \end{array} \right.$ es obviamente incompatible (¡justificar!).

Si se lo escribe en la forma (2.22), se obtiene $x \begin{pmatrix} 1 \\ 2 \end{pmatrix} + y \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, es decir,

$(x + 2y) \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, lo cual es imposible. Toda combinación lineal de las columnas de $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ es proporcional a $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ y por lo tanto, nunca puede ser igual a $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Problema 2.4.2
Escribir los siguientes sistemas en las formas matriciales (2.18) y (2.22). Resolverlos y verificar en el caso compatible que las soluciones satisfacen las ecuaciones matriciales.

a) $\left\{ \begin{array}{rcl} x + y + z & = & 1 \\ 2x + y + z & = & 2 \end{array} \right.$ b) $\left\{ \begin{array}{rcl} x + y & = & 5 \\ x - y & = & 1 \\ 5x + y & = & 17 \end{array} \right.$

2.4.1. Sistemas homogéneos y vectores ortogonales

Dos vectores $\boldsymbol{a}$, $\boldsymbol{b}$ de n componentes son ortogonales (es decir, "perpendiculares") si su producto escalar es nulo:

$$\boldsymbol{a}\cdot\boldsymbol{b} = \begin{pmatrix} a_1 & \dots & a_n \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + \ldots + a_n b_n = 0$$

Consideremos ahora un sistema lineal homogéneo de $m \times n$

$$\left\{ \begin{array}{ccc} a_{11}x_1 + \ldots + a_{1n}x_n & = & 0 \\ \vdots & \vdots & \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n & = & 0 \end{array} \right. \tag{2.25}$$

es decir, en forma matricial,

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \tag{2.26}$$

o en forma compacta,

$$A\mathbf{x} = \mathbf{0} \tag{2.27}$$

Tanto (2.25) como (2.26) o (2.27) muestran que toda solución $\mathbf{x}$ del sistema lineal homogéneo anterior es necesariamente un vector **ortogonal a todas las filas** $\boldsymbol{a}_{i*} = (a_{i1}, \ldots, a_{in})$ de A, ya que la ecuación iésima en (2.25) es justamente

$$\boldsymbol{a}_{i*}\cdot\mathbf{x} = 0$$

Si el sistema es **compatible determinado**, el único vector ortogonal a todas las filas es $\mathbf{x} = \mathbf{0}$.
Pero si el sistema es **compatible indeterminado**, existen infinitos vectores no nulos $\mathbf{x}$ **ortogonales a todas las filas** de A.

El significado geométrico de resolver un sistema lineal homogéneo es encontrar todos los vectores $\mathbf{x}$ **ortogonales** a todas las filas de A.

Ejemplo 2.4.3
El sistema $\left\{ \begin{array}{ccc} x - 2y & = & 0 \\ -3x + 6y & = & 0 \end{array} \right.$, es decir $\begin{pmatrix} 1 & -2 \\ -3 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ tiene como solución el conjunto $\{y \begin{pmatrix} 2 \\ 1 \end{pmatrix}, y \in \mathbb{R}\}$, siendo todo vector $y \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ ortogonal a los dos filas de $\begin{pmatrix} 1 & -2 \\ -3 & 6 \end{pmatrix}$, que son proporcionales.

Problema 2.4.3
Encuentre los vectores ortogonales a todas las filas de las matrices
a) $A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, b) $B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix}$

2.5. Matriz inversa

Una de las ventajas que ofrece el formalismo matricial es el de poder resolver los sistemas lineales de $n \times n$ compatibles determinados en forma análoga a como se resuelve un sistema $ax = b$ de 1×1 cuando $a \neq 0$. En primer lugar, notamos que la matriz identidad definida en (2.6), de elementos $(I)_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$,

$$I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

es el **elemento neutro** para la multiplicación de matrices:

Producto por matriz identidad.
Si I_n, I_m son las matrices identidad de $n \times n$ y $m \times m$, entonces $\forall$ A de $m \times n$,

$$A\, I_n = A\,, \quad I_m\, A = A$$

Demostración: $(AI_n)_{ij} = \sum_{k=1}^{n} a_{ik}(I_n)_{kj} = a_{ij}(I_n)_{jj} = a_{ij}$ pues $(I_n)_{kj} = 0$ si $k \neq j$ y $(I_n)_{jj} = 1$. Por lo tanto, $AI_n = A$. La demostración de $I_m\, A = A$ es similar.
Por ejemplo,

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

En particular, si A es de $n \times n$, entonces

$$A\, I_n = I_n\, A = A$$

Por ejemplo,

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

Dada una matriz A de $n \times n$, podemos preguntarnos ahora si existe una **matriz inversa** A^{-1} tal que $A\, A^{-1} = A^{-1}A = I_n$. En lo que sigue I denota la matriz I_n.

Definición.
Una matriz A de dimensión $n \times n$ se dice *no-singular* o *invertible* cuando existe una matriz A^{-1} de dimensión $n \times n$ tal que

$$A\, A^{-1} = A^{-1}\, A = I$$

La matriz A^{-1} se llama “inversa multiplicativa” o inversa de A.

Surge una pregunta natural: de existir, ¿es la inversa única?

Si A de $n \times n$ es invertible, **la inversa A^{-1} es única**.
Demostración. Si existiesen B y C de $n \times n$ ambas matrices inversas de A entonces

$$B = B\,I = B\,(A\,C) = (B\,A)\,C = I\,C = C$$

por lo que $B = C$. La única inversa de A se la denota entonces A^{-1}.

Observación: La demostración anterior muestra que si A tiene una inversa a izquierda B y una inversa a derecha C, necesariamente son coincidentes y únicas. Esto puede ocurrir sólo cuando A es cuadrada, en cuyo caso si tiene una inversa a izquierda, también tiene inversa a derecha y viceversa. En cambio, las matrices no cuadradas de $m \times n$ pueden tener una inversa a derecha (si $m < n$) o a izquierda (si $m > n$) pero no ambas simultáneamente. Además, si existen no son únicas. No estudiaremos el caso $m \neq n$ aquí.

Definición
Una matriz de dimensión $n \times n$ se dice *singular* o *no-invertible* si no tiene matriz inversa.

Ejemplo 2.5.1

Si $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ se observa que para cualquier matriz B de 2×2,

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{11} \\ b_{21} & b_{21} \end{pmatrix} \neq I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

pués b_{11} (y también b_{21}) tendría que ser igual a 1 y a 0 al mismo tiempo. Entonces no existe B que sea inversa de A. Luego A es una matriz singular (no tiene inversa).

Ejemplo 2.5.2

Si $A = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}$, planteando

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 2b_{11} - b_{12} & b_{11} \\ 2b_{21} - b_{22} & b_{21} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

se obtiene un sistema para los b_{ij} cuya única solución es (¡probar!)
$b_{11} = 0$, $b_{12} = -1$, $b_{21} = 1$, $b_{22} = 2$:

$$\begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix}$$

Por lo tanto, esta matriz A es **no singular** o **invertible** y su inversa es $A^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix}$.

Conclusión 1
Las matrices cuadradas ($n \times n$) pueden dividirse en dos clases:

- no-singulares (invertibles)
- singulares (no-invertibles)

Cada una de estas clases tiene ciertas propiedades que serán enunciadas y exploradas en este y los siguientes capítulos.
Por otra parte, las matrices no cuadradas ($m \times n$, con $m \neq n$) no pueden ser clasificadas o categorizadas en una forma tan simple como en el caso $m = n$.

Conclusión 2
Si A de $n \times n$ es **no singular** (invertible), el sistema de $n \times n$

$$A\mathbf{x} = \mathbf{b} \tag{2.28}$$

es **compatible determinado** $\forall\ \boldsymbol{b} \in \mathbb{R}^n$, siendo la única solución

$$\mathbf{x} = A^{-1}\boldsymbol{b}$$

Demostración.
En primer lugar, el sistema es compatible pues $\mathbf{x} = A^{-1}\boldsymbol{b}$ es solución del sistema:

$$A\mathbf{x} = A(A^{-1}\boldsymbol{b}) = (AA^{-1})\boldsymbol{b} = I\boldsymbol{b} = \boldsymbol{b}$$

Y es compatible determinado (solución única) pues si existe algun $\mathbf{x}$ que satisface (2.28), multiplicando ambos miembros a izquierda por A^{-1} obtenemos

$$A^{-1}(A\mathbf{x}) = (A^{-1}A)\mathbf{x} = I\mathbf{x} = \mathbf{x} = A^{-1}\boldsymbol{b}$$

por lo que necesariamente $\mathbf{x} = A^{-1}\boldsymbol{b}$.

En particular, el sistema homogéneo $A\mathbf{x} = \mathbf{0}$ tiene sólo la solución trivial: $\mathbf{x} = A^{-1}\mathbf{0} = \mathbf{0}$.

La conclusión anterior también implica obviamente que si el sistema de $n \times n$ $A\mathbf{x} = \boldsymbol{b}$ es incompatible o compatible indeterminado (infinitas soluciones), A es singular, pues de lo contrario sería compatible determinado.

Mostraremos luego que la reciproca es también válida: si el sistema cuadrado es compatible determinado para algún $\boldsymbol{b} \in \mathbb{R}^n$, entonces necesariamente A es no singular y por lo tanto compatible determinado $\forall\ \boldsymbol{b}$.

Ejemplo 2.5.3 El sistema $\begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ es compatible determinado ya que $A = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}$ es invertible (ejemplo 2.6.2). La solución es entonces

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A^{-1}\begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} -3 \\ 7 \end{pmatrix}$$

o sea, $x_1 = -3$, $x_2 = 7$.

2.5.1. Reglas para matrices inversas

Sea A una matriz de $n \times n$ no-singular (invertible). Entonces

1. A^{-1} es no-singular y $(A^{-1})^{-1} = A$.
   Demostración.
   Obviamente, como $A\,A^{-1} = A^{-1}\,A = I$, A^{-1} es invertible, siendo A su inversa.

2. Si $\alpha \neq 0$, $(\alpha\,A)$ es no-singular y $(\alpha A)^{-1} = \frac{1}{\alpha}\,A^{-1}$.
   Demostración.
   $(\alpha A)\left(\frac{1}{\alpha}A^{-1}\right) = \left(\alpha\frac{1}{\alpha}\right)(AA^{-1}) = 1\,I = I$. De igual forma se prueba $\left(\frac{1}{\alpha}A^{-1}\right)(\alpha A) = I$

3. Si B es también de $n \times n$ y no singular, entonces AB es no singular y

   $$(A\,B)^{-1} = B^{-1}\,A^{-1}$$

   Notar **la inversión del orden** en el producto.
   Demostración.
   Utilizando la asociatividad del producto,

   $$(A\,B)(B^{-1}A^{-1}) = (A(B\,B^{-1}))A^{-1} = (A\,I)A^{-1} = A\,A^{-1} = I$$

   De igual forma se prueba que $(B^{-1}\,A^{-1})(AB) = I$.

4. El resultado anterior se puede extender a varias matrices de $n \times n$ no-singulares $A_1, A_2, \ldots, A_k$: El producto $A_1\,A_2 \ldots A_k$ es no singular y su inversa es

   $$(A_1\,A_2 \ldots A_k)^{-1} = A_k^{-1} \ldots A_2^{-1} A_1^{-1}$$

   Se deja la demostración para el lector. Por ejemplo, si A, B y C son todas de $n \times n$ y no singulares,

   $$(ABC)^{-1} = C^{-1}\,B^{-1}\,A^{-1}$$

5. Si A es no singular $\Rightarrow$ A^T es no singular y su inversa es la traspuesta de A^{-1}:

   $$(A^T)^{-1} = (A^{-1})^T$$

   Demostración.
   Partiendo de $AA^{-1} = I$ y trasponiendo ambos miembros se obtiene

   $$(AA^{-1})^T = (A^{-1})^T A^T = I^T = I$$

   En forma análoga, partiendo de $A^{-1}A = I$ se obtiene $A^T(A^{-1})^T = I$. Por lo tanto $(A^T)^{-1} = (A^{-1})^T$.

Ejemplo 2.5.4: El caso de 2×2.
La matriz
$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$
es invertible si y sólo si
$$ad - bc \neq 0$$
En tal caso
$$A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \quad (2.29)$$
Esto será mostrado en forma general en el próximo capítulo. Por el momento, podemos fácilmente verificar que
$$\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} da-bc & db-bd \\ -ca+ac & -cb+ad \end{pmatrix} = \begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix} = (ad-bc)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
por lo que si $ad - bc \neq 0$,
$$\frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
que implica (2.29). Análogamente se prueba que $AA^{-1} = I_2$. Es fácil verificar que si $ad - bc = 0$ no existe inversa (se deja su demostración para el lector).
Por ejemplo, retornando a los ejemplos 2.6.1 y 2.6.2, vemos que $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ no es invertible pues $ad - bc = 0$, mientras que $\begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}$ sí lo es pues $ad - bc = 1$. La inversa obtenida en 2.6.2, $A^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix}$ puede también obtenerse de (2.29).

El número "mágico" $ad - bc$ que decide si A es invertible o no se denomina *determinante*. Su expresión para matrices de $n \times n$ es más compleja y la daremos en el próximo capítulo.

Problemas 2.5.1

1. Probar que si A es de $n \times n$ y no singular y $AB = AC$, con B, C de $n \times p \Rightarrow B = C$. ¿Puede afirmarse lo mismo si A es singular?
2. Probar que si A de $n \times n$ es no singular, $(A^2)^{-1} = (A^{-1})^2$ y en general, $(A^k)^{-1} = (A^{-1})^k$ para $k \geq 1$ (natural).
3. Sean A, B de $n \times n$ no singulares. Expresar en términos de A, A^{-1} y B^{-1}, i) $(AB^2)^{-1}$ ii) $(ABA^{-1})^{-1}$, iii) $((AB)^T)^{-1}$
4. Probar que si $ad - bc = 0$ entonces $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ no tiene inversa.
5. Resolver el sistema $\begin{pmatrix} 1 & 3 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$, determinando la inversa (en caso de que exista) de la matriz de coeficientes.

2.5.2. Inversa de matrices ortogonales

Un caso especial de matrices de inversión sencilla es el de las matrices **ortogonales**. Una matriz A de $n \times n$ se dice ortogonal (u ortonormal) si todas sus columnas $\boldsymbol{a}_{*i}$ son ortogonales entre si y de longitud 1, es decir, si forman un **conjunto ortonormal**:

$$\boldsymbol{a}_{*i} \cdot \boldsymbol{a}_{*j} = a_{1i}a_{1j} + \ldots + a_{ni}a_{nj} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

es decir, $\boldsymbol{a}_{*i} \cdot \boldsymbol{a}_{*j} = I_{ij}$, donde I_{ij} denota el elemento i, j de la matriz identidad.

Por ejemplo,

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \tag{2.30}$$

es una matriz ortogonal (¡comprobar!). Para la matriz (2.30) se comprueba que las filas forman también un conjunto ortonormal y que $A^{-1} = A^T$:

$$A^T A = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = AA^T$$

Estos resultados son válidos para **toda** matriz ortogonal:

Si A es una matriz de $n \times n$ ortogonal $\Rightarrow$ es **no singular** y su inversa es su **traspuesta**:

$$A^{-1} = A^T \quad \text{es decir,} \quad A^T A = AA^T = I$$

Asimismo, si $A^{-1} = A^T \Rightarrow A$ es una matriz ortogonal.

Demostración: Como la fila i de A^T es la columna i de A, si A es ortogonal obtenemos, a partir de la definción de producto matricial,

$$(A^T A)_{ij} = \boldsymbol{a}_{*i} \cdot \boldsymbol{a}_{*j} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} = I_{ij} \tag{2.31}$$

por lo que $A^T A = I$. Para A cuadrada esto implica $AA^T = I$ y $A^{-1} = A^T$.
Y si $A^{-1} = A^T \Rightarrow A^T A = I$, es decir, $(A^T A)_{ij} = \boldsymbol{a}_{i*} \cdot \boldsymbol{a}_{j*} = I_{ij}$, lo que implica que las columnas de A son ortonormales, es decir, que A es ortogonal.

Si A de $n \times n$ es ortogonal $\Rightarrow A^T$ es también ortogonal, ya que $(A^T)^{-1} = A = (A^T)^T$.
Esto implica que **las filas de A forman también un conjunto ortonormal**, ya que son las columnas de A^T. Esto puede también obtenerse de $AA^T = I$.
En resumen, si las columnas de una matriz A de $n \times n$ forman un conjunto ortonormal, también lo forman las filas, y viceversa.

Problemas 2.5.2

1. Verificar que las siguientes matrices son ortogonales y obtener su inversa:

a) $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ b) $B = \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 & \sqrt{\frac{2}{3}} \\ 0 & -1 & 0 \\ -\sqrt{\frac{2}{3}} & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}$ c) $C = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix}$

2. Muestre que si $A^T A = AA^T = I \Rightarrow$ las columnas y las filas de A son conjuntos ortonormales.

2.6. Matrices elementales y sistemas lineales

El objetivo es resolver el sistema lineal $A\mathbf{x} = \boldsymbol{b}$ usando un sistema modificado, equivalente al original, mediante sucesivas multiplicaciones por matrices simples que representan las operaciones por filas.

2.6.1. Sistemas equivalentes

Consideremos un sistema lineal

$$A\mathbf{x} = \boldsymbol{b} \tag{2.32}$$

compatible, de dimensión $m \times n$. Si se multiplica ambos lados del sistema anterior por una matriz M no-singular (invertible) de $m \times m$ se obtiene

$$MA\mathbf{x} = M\boldsymbol{b} \tag{2.33}$$

Entonces:

- Si $\mathbf{x}$ es solución del sistema (2.32) $\Rightarrow$ también satisface el sistema (2.33). Es decir, toda solución del primero es también solución del segundo sistema.
- A su vez, si $\mathbf{x}$ es una solución del sistema (2.33) también es solución del sistema (2.32), ya que si se multiplicase ambos lados de (2.33) por M^{-1} se obtendría

$$\begin{aligned} M^{-1}(MA\mathbf{x}) &= M^{-1}(M\boldsymbol{b}) \\ \left(M^{-1}M\right)A\mathbf{x} &= \left(M^{-1}M\right)\boldsymbol{b}, \end{aligned}$$

  resultando $A\mathbf{x} = \boldsymbol{b}$ ya que $M^{-1}M = I$.

Así hemos demostrado que *los dos sistemas son equivalentes*, siempre y cuando la matriz M sea invertible.

Para obtener un sistema equivalente a $A\mathbf{x} = \boldsymbol{b}$ pero más fácil de resolver, se multiplicarán ambos lados del sistema por una sucesión de matrices *no-singulares* $E_1, E_2, \ldots, E_k$ de $m \times m$, de modo de llegar a un sistema más sencillo, es decir:

$$E_k \ldots E_2 E_1\, A\,\mathbf{x} = E_k \ldots E_2 E_1\, \boldsymbol{b}$$

Si denominamos $M = E_k \ldots E_2 E_1$, entonces

$$MA\,\mathbf{x} = M\mathbf{b}$$

Esto transforma el sistema original en un sistema más sencillo

$$U\mathbf{x} = \boldsymbol{c}$$

donde

$$\begin{aligned} U &= MA = E_k \ldots E_2 E_1 A \\ \mathbf{c} &= M\boldsymbol{b} = E_k \ldots E_2 E_1\, \boldsymbol{b} \end{aligned}$$

Conclusión
Como todas las matrices E_i son no-singulares, el producto

$$M = E_k \dots E_2 E_1$$

también es no-singular.
Por lo tanto, el sistema $A\mathbf{x} = \boldsymbol{b}$ y el resultante $U\mathbf{x} = \boldsymbol{c}$ son equivalentes.

2.6.2. Matrices elementales

Definición
Una *matriz elemental* es una matriz cuadrada $m \times m$ que se obtiene a partir de la matriz identidad I_m mediante una operación elemental sobre sus filas.

Como existen tres tipos de operaciones elementales sobre las filas, existen tres tipos de matrices elementales:

Tipo I: Intercambiar dos filas de I_m.

Por ejemplo, si $m = 3$, el intercambio de la fila 2 por la fila 3 implica

$$E_I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

Si A es una matriz de 3×3, multiplicando por E_I resulta

$$E_I A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

Por otra parte, si se multiplica A por la derecha por E_I, resulta

$$AE_I = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{13} & a_{12} \\ a_{21} & a_{23} & a_{22} \\ a_{31} & a_{33} & a_{32} \end{pmatrix}$$

Por lo tanto:
Multiplicar a izquierda por la matriz E_I intercambia las **filas** 2 y 3 de la matriz A.
En cambio, multiplicar a derecha, intercambia las **columnas** 2 y 3 de la matriz A.

Tipo II: Multiplicar una fila de I_m por un escalar no nulo.

Por ejemplo, si se multiplica la fila 3 por el número α,

$$E_{II} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \alpha \end{pmatrix}$$

Si A es cualquier matriz de 3×3, entonces

$$E_{II}A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \alpha \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ \alpha a_{31} & \alpha a_{32} & \alpha a_{33} \end{pmatrix}$$

$$AE_{II} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \alpha \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \alpha a_{13} \\ a_{21} & a_{22} & \alpha a_{23} \\ a_{31} & a_{32} & \alpha a_{33} \end{pmatrix}$$

Por lo tanto:
Multiplicar a izquierda por la matriz E_{II} multiplica la **fila** 3 de la matriz A por α.
En cambio, multiplicar a derecha multiplica la **columna** 3 por el escalar α.

Tipo III: Sumar a una fila de I_m, algún múltiplo no nulo de otra fila de I_m.

Ejemplo: Se suma a la fila 3, la fila 1 multiplicada por α.

$$E_{III} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \alpha & 0 & 1 \end{pmatrix}$$

Si A es cualquier matriz de 3×3, entonces

$$E_{III}A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \alpha & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} + \alpha a_{11} & a_{32} + \alpha a_{12} & a_{33} + \alpha a_{33} \end{pmatrix}$$

$$AE_{III} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \alpha & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} + \alpha a_{13} & a_{12} & a_{13} \\ a_{21} + \alpha a_{23} & a_{22} & a_{23} \\ a_{31} + \alpha a_{33} & a_{32} & a_{33} \end{pmatrix}$$

Por lo tanto:
Multiplicar a izquierda por la matriz E_{III} suma a la **fila** 3 de A, α veces la fila 1.
En cambio, multiplicar a derecha suma a la **columna** 1, α veces la columna 3.

Lema 2.6.1: Si E es una matriz elemental (de Tipo I, II o III) entonces E es no-singular y E^{-1} es una matriz elemental del mismo tipo.

Demostración. Separamos, según sea E una matriz elemental de Tipo I, II o III.

Tipo I: Si E_I intercambia dos filas de la matriz identidad, entonces E_I puede volver atrás el cambio intercambiando de nuevo las filas. Por lo tanto

$$E_I E_I = I$$

por lo que E_I es invertible y $E_I^{-1} = E_I$ (la inversa es del mismo tipo).

Tipo II: Si E_{II} se forma al multiplicar alguna fila de I por un escalar $\alpha \neq 0$ entoces podemos proponer como matriz inversa, aquella que multiplique la misma fila de la matriz I por el escalar $1/\alpha$. Por lo tanto E_{II} es invertible y su inversa es del mismo tipo.

Tipo III: Si E_{III} se forma al sumar "α veces la fila j a la fila i"de la matriz I

$$E_{III} = \begin{pmatrix} 1 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & & & & & \vdots \\ 0 & \cdots & 1 & \cdots & \cdots & \cdots & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & \cdots & \alpha & \cdots & 1 & \cdots & 0 \\ \vdots & & & & & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix} \begin{matrix} \\ \\ \text{fila } j \\ \\ \text{fila } i \\ \\ \\ \end{matrix}$$

la matriz inversa es aquella que resta α veces la fila j a la fila i de la matriz I:

$$E_{III}^{-1} = \begin{pmatrix} 1 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & & & & & \vdots \\ 0 & \cdots & 1 & \cdots & \cdots & \cdots & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & \cdots & -\alpha & \cdots & 1 & \cdots & 0 \\ \vdots & & & & & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix} \begin{matrix} \\ \\ \text{fila } j \\ \\ \text{fila } i \\ \\ \\ \end{matrix}$$

Se verifica que $E_{III}^{-1}\, E_{III} = E_{III}\, E_{III}^{-1} = I$.

■

El proceso de reducción por filas de una matriz (eliminación Gaussiana) es pues equivalente a realizar sucesivas multiplicaciones por matrices elementales E_i.

Definición
Una matriz cuadrada B de $n \times n$ es *equivalente por filas* a otra matriz cuadrada A de $n \times n$ si existe una cantidad finita de matrices elementales $E_1, E_2, \ldots, E_k$ tales que

$$B = E_k \ldots E_2 E_1\, A$$

Es decir, B es "equivalente"por filas a A si B se puede obtener a partir de A mediante una cantidad finita de operaciones elementales sobre filas.

Comentario. Dos resultados obvios:

1. Si B es equivalente por filas a A, entonces A es equivalente por filas con B.
2. Si B es equivalente por filas a A, y A es equivalente por filas a C, entonces B es equivalente por filas a C.

Teorema 2.6.1 Si A es una matriz cuadrada de $n \times n$, entonces las siguientes proposiciones son equivalentes:

i) A es no-singular (tiene inversa)

ii) El sistema lineal $A.\mathbf{x} = 0$ tiene solamente la solución trivial ($\mathbf{x} = 0$).

iii) A es equivalente por filas a la matriz identidad I_n de $n \times n$.

Demostración.

(i) → (ii) Esto ya fue demostrado en 2.6 (Ec. (2.28)). Si $\mathbf{x}$ es solución del sistema homogéneo, multiplicando por A^{-1} ambos lados de $A.\mathbf{x} = \mathbf{0}$,

$$A^{-1}\,A\mathbf{x} = A^{-1}\,\mathbf{0}$$

se obtiene

$$\mathbf{x} = \mathbf{0}$$

(ii) → (iii) Utilizando operaciones elementales sobre las filas se transforma el sistema $A.\mathbf{x} = \mathbf{0}$ en otro sistema $U.\mathbf{x} = \mathbf{0}$, donde U es una matriz escalonada reducida de Gauss- Jordan. Si U no fuese la identidad, alguno de los coeficientes de la diagonal principal de U sería cero. Eso significaría que A no es equivalente por filas con I, y entonces la última fila de U debe tener todos sus elementos nulos. Tal característica dice que $A\mathbf{x} = \mathbf{0}$ es equivalente a un sistema homogéneo con más incógnitas que ecuaciones. Eso diría que el sistema $A\mathbf{x} = \mathbf{0}$ debe tener "infinitas soluciones no triviales", lo cual es contradictorio respecto de (ii). Por tanto, U tiene que ser la matriz identidad.

(iii) → (i) Como A es equivalente por filas con I, entonces existe una número finito de matrices elementales, no singulares, $E_1, E_2, \ldots, E_k$ tales que

$$E_k \ldots E_2 E_1 A = I$$

Luego multiplicando por la inversa de $E_k \ldots E_2\, E_1$, se obtiene

$$A = (E_k \ldots E_2 E_1)^{-1} I$$

Por lo tanto A es no-singular, por ser producto de matrices no-singulares, e igual a

$$A = E_1^{-1} E_2^{-1} \ldots E_k^{-1}$$

■

Observación: El paso (**i**) → (**ii**) sólo asume en principio existencia de la inversa a izquierda de A. Como **(ii)** implica **(iii)** para A de $n \times n$, el sistema $A\mathbf{x} = \boldsymbol{b}$ tendrá solución única $\forall\ \boldsymbol{b} \in \mathbb{R}^n$, en particular para los vectores columna $\boldsymbol{b}$ de I_n. Por lo tanto $AB = I_n$ tiene también solución, es decir, A tendrá inversa a derecha, la cual necesariamente coincidirá con la inversa a izquierda como se demostró en 2.5.

Corolario. (Importante)
Un sistema lineal $A\mathbf{x} = \mathbf{b}$ de $n \times n$ tiene solución única $\Leftrightarrow$ A es no-singular (invertible).

Demostración

$\Leftarrow$ Si A es no-singular hemos ya demostrado en 2.28 que el sistema $A\mathbf{x} = \boldsymbol{b}$ es compatible con solución única

$$\mathbf{x} = A^{-1}\boldsymbol{b}$$

También se puede deducir esto a partir de la equivalencia con (iii) del teorema anterior.

$\Rightarrow$ Sea $\mathbf{x}_1$ la solución única del sistema $A\mathbf{x} = \boldsymbol{b}$. Si A fuese singular, entonces por el teorema anterior y la equivalencia entre (i) e (ii) tendríamos que el sistema homogéneo $A\mathbf{x} = \mathbf{0}$ no tendría solución única. Así $A\mathbf{x} = \mathbf{0}$ tendría tendría infinitas soluciones no triviales, por ejemplo una solución $\mathbf{z} \neq \mathbf{0}$. En tal caso, si $\mathbf{x}_2 = \mathbf{x}_1 + \mathbf{z}$, tenemos $\mathbf{x}_2 \neq \mathbf{x}_1$ y

$$A\mathbf{x}_2 = A(\mathbf{x}_1 + \mathbf{z}) = A\mathbf{x}_1 + A\mathbf{z} = \boldsymbol{b} + \mathbf{0} = \boldsymbol{b}$$

por lo que $\mathbf{x}_2$ sería otra solución del sistema $A\mathbf{x} = \boldsymbol{b}$, distinta de $\mathbf{x}_1$. Esta conclusión es absurda, ya que por hipótesis $\mathbf{x}_1$ es la única solución.

Esto muestra que A no puede ser singular. Si $A\mathbf{x} = \boldsymbol{b}$ tiene una sóla solución entonces A es no-singular.

$\blacksquare$

Síntesis
Hasta el momento tenemos como resumen las siguientes equivalencias: Si A es $n \times n$,

- A es no-singular (invertible).
- A es equivalente por filas a la matriz identidad.
- El sistema lineal $A\mathbf{x} = 0$ tiene solución única (la solución trivial).
- El sistema lineal $A\mathbf{x} = \mathbf{b}$ tiene solución única ($\mathbf{x} = A^{-1}\mathbf{b}$) $\forall\, \boldsymbol{b} \in \mathbb{R}^n$

Problemas 2.5.1

1. Una es no-singular mientras que la otra es singular. Analizar, y decidir.

(a) $\begin{pmatrix} 1 & 3 \\ 4 & -12 \end{pmatrix}$ **(b)** $\begin{pmatrix} 1 & 3 \\ 4 & 12 \end{pmatrix}$

2. ¿Singular o no-singular?

(a) $\begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$ **(b)** $\begin{pmatrix} 1 & 2 \\ -3 & -6 \end{pmatrix}$ **(c)** $\begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 4 & 1 \end{pmatrix}$ **(d)** $\begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 3 \\ 3 & 4 & 7 \end{pmatrix}$

3. Describir las matrices que son equivalentes a

(a) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ **(b)** $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ **(c)** $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ **(d)** $\begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}$

4. ¿Pueden dos matrices equivalentes tener diferente dimensión?

5. Dar dos matrices escalonadas reducidas que tengan sus coeficientes pivotes en la misma columna pero que no sean equivalentes.

6. Extender la definición de equivalencia por filas de matrices a sistemas equivalentes por filas.

7. Probar que cualquier sistema lineal con una matriz de coeficientes no-singular tiene solución y es única.

2.7. Método para determinar la matriz inversa

Si A es no-singular, entonces A es equivalente por filas a la matriz identidad I. Esto es, mediante matrices elementales adecuadas se obtiene

$$E_k \dots E_2 E_1 A = I \tag{2.34}$$

Esto implica que

$$E_k \dots E_2 E_1 = A^{-1} \tag{2.35}$$

(multiplicando ambos miembros de (2.34) a derecha por A^{-1}, se obtiene $E_k \dots E_2 E_1 A A^{-1} = I A^{-1}$ y por lo tanto, (2.35)).

Conclusión.
La misma sucesión de operaciones elementales por filas que transforman la matriz A no-singular en la matriz identidad, también transforman la matriz identidad en la matriz inversa A^{-1}, ya que $A^{-1} = E_k \dots E_2\, E_1 = E_k \dots E_2\, E_1\, I$.

Por lo tanto, **el procedimiento práctico para determinar la inversa A^{-1} es:**

(i) Se forma la matriz aumentada

$$(A \mid I)$$

(de dimensión $n \times 2n$).

(ii) Se aplican las operaciones elementales para llevar a A a la forma escalonada reducida de Gauss-Jordan, resultando

$$\left(I \mid A^{-1}\right)$$

o sea,

$$E_k \dots E_2 E_1 \,(A \mid I) = \left(I \mid A^{-1}\right)$$

de donde se obtiene la inversa A^{-1}.

Notemos también que si A es no-singular y $\mathbf{x}$ es la única solución del sistema $A\mathbf{x} = \mathbf{b}$, entonces la forma escalonada de Gauss-Jordan de la matriz aumentada $(A \mid \mathbf{b})$, de dimensión $n \times (n+1)$, es necesariamente

$$\left(I \mid A^{-1}\boldsymbol{b}\right)$$

dado que A es equivalente por filas a la identidad y el lado derecho debe ser la solución única del problema.

Ejemplo 2.7.1

Dada $A = \begin{pmatrix} 1 & 4 & 3 \\ -1 & -2 & 0 \\ 2 & 2 & 3 \end{pmatrix}$, para hallar A^{-1} se realiza la reducción por filas hasta llegar a la forma de Gauss-Jordan:

$$\begin{aligned}
(A \mid I) \quad &= \quad \left(\begin{array}{ccc|ccc} 1 & 4 & 3 & 1 & 0 & 0 \\ -1 & -2 & 0 & 0 & 1 & 0 \\ 2 & 2 & 3 & 0 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 4 & 3 & 1 & 0 & 0 \\ 0 & 2 & 3 & 1 & 1 & 0 \\ 0 & -6 & -3 & -2 & 0 & 1 \end{array}\right) \\
&\longrightarrow \left(\begin{array}{ccc|ccc} 1 & 4 & 0 & \frac{1}{2} & -\frac{3}{2} & -\frac{1}{2} \\ 0 & 2 & 0 & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 6 & 1 & 3 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ 0 & 0 & 1 & \frac{1}{6} & \frac{1}{2} & \frac{1}{6} \end{array}\right) \\
\Rightarrow \; A^{-1} \quad &= \quad \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{6} \end{pmatrix}
\end{aligned}$$

Por ejemplo, la solución del sistema lineal

$$A\,\mathbf{x} = \begin{pmatrix} 12 \\ -12 \\ 8 \end{pmatrix}$$

es

$$\mathbf{x} = A^{-1}\boldsymbol{b} = \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{6} \end{pmatrix} \begin{pmatrix} 12 \\ -12 \\ 8 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \\ -\frac{8}{3} \end{pmatrix}$$

Problemas 2.7

1. Determinar si las siguientes matrices son no singulares. En tal caso, hallar su inversa mediante reducción de la matriz ampliada $(A|I_n)$.

$$(a)\ \begin{pmatrix} 3 & 4 \\ 2 & -1 \end{pmatrix} \qquad (b)\ \begin{pmatrix} 3 & 2 & 1 \\ 6 & -4 & 0 \\ 0 & 1 & 1 \end{pmatrix} \qquad (c)\ \begin{pmatrix} 1 & 2 & -1 \\ 2 & 4 & 0 \\ 0 & 1 & -3 \end{pmatrix} \qquad (d)\ \begin{pmatrix} 1 & 0 & 2 \\ -1 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$$

2. En el caso (d) anterior, utilizar la inversa para hallar la solución del sistema $A\mathbf{x} = \boldsymbol{b}$, con $\boldsymbol{b} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$.

3. En el caso (a) anterior, escriba en orden las matrices elementales correspondientes a las operaciones necesarias para llevar A a la forma reducida escalonada de Gauss-Jordan y exprese A^{-1} como producto de las mismas.

4. Resolver cada sistema utilizando notación matricial y expresar la solución utilizando vectores. Aclarar si la matriz A del sistema tiene inversa o no y hallar la inversa si existe.

   (a) $\begin{aligned} 3x + 6y &= 18 \\ x + 2y &= 6 \end{aligned}$ **(b)** $\begin{aligned} x + y &= 1 \\ x - y &= -1 \end{aligned}$ **(c)** $\begin{aligned} x_1 \quad\quad + x_3 &= 4 \\ x_1 - x_2 + 2x_3 &= 5 \\ 4x_1 - x_2 + 5x_3 &= 17 \end{aligned}$

   (d) $\begin{aligned} 2a + b - c &= 2 \\ 2a \quad + c &= 3 \\ a - b \quad &= 0 \end{aligned}$ **(e)** $\begin{aligned} x \quad + z + w &= 4 \\ 2x + y \quad - w &= 2 \\ 3x + y + z \quad &= 7 \end{aligned}$

5. Muestre que si A de $(n+m) \times (n+m)$ es una matriz definida por bloques de la forma
$$A = \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix}$$
con B de $n \times n$ y C de $m \times m$, entonces A es no singular si y sólo si B y C son no singulares, con
$$A^{-1} = \begin{pmatrix} B^{-1} & 0 \\ 0 & C^{-1} \end{pmatrix}$$

6. Utilizando el resultado anterior, determine la inversa de
$$A = \begin{pmatrix} 1 & 3 & 0 & 0 \\ 1 & 4 & 0 & 0 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 2 & 3 \end{pmatrix}$$

7. Resuelva los problemas 1, 2 y 4 utilizando un software adecucado para el manejo de matrices.

2.8. Factorización triangular (LU)

Si una matriz A de $n \times n$ no singular puede reducirse a una forma triangular superior U sólo usando operaciones por filas de tipo III, entonces es posible expresar el proceso de reducción mediante una factorización matricial
$$A = LU$$
donde L es triangular inferior y de la forma
$$L = \begin{pmatrix} 1 & 0 & \dots & 0 \\ l_{21} & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ l_{n1} & l_{n2} & \dots & 1 \end{pmatrix}$$

En efecto, en tal caso $E_k \dots E_1 A = U$, por lo que

$$A = (E_k \dots E_1)^{-1} U = E_1^{-1} \, \dots \, E_k^{-1} U$$

que implica

$$L = E_1^{-1} \, \dots \, E_k^{-1}$$

Pero al ser todas las E_i operaciones sucesivas de tipo III, L es triangular inferior, con $L_{ii} = 1 \; \forall \; i$. Por ejemplo, si para A de 3×3 E_1 realiza la operación $f_2 - \alpha_1 f_1$, E_2 la operación $f_3 - \alpha_2 f_1$ y E_3 la operación $f_3 - \alpha_3 f_2$, entonces

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -\alpha_1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \; E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\alpha_2 & 0 & 1 \end{pmatrix}, \; E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\alpha_3 & 1 \end{pmatrix}$$

y

$$E_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ \alpha_1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \; E_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \alpha_2 & 0 & 1 \end{pmatrix}, \; E_3^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \alpha_3 & 1 \end{pmatrix}$$

son también operaciones de tipo III (con $\alpha_i \to -\alpha_i$). Por lo tanto,

$$L = E_1^{-1} E_2^{-1} E_3^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ \alpha_1 & 1 & 0 \\ \alpha_2 & \alpha_3 & 1 \end{pmatrix} \tag{2.36}$$

La descomposición LU es un resultado útil en la resolución numérica de sistemas: un método eficiente para resolver sistemas grandes $A\mathbf{x} = \boldsymbol{b}$ es precisamente escribirlo como

$$LU\mathbf{x} = \boldsymbol{b}$$

y resolverlo en dos pasos:
I) Resolver $L\mathbf{y} = \boldsymbol{b}$, mediante sustitución hacia adelante (por ser L triangular inferior)
II) Resolver $U\mathbf{x} = \mathbf{y}$, con $\mathbf{y}$ el valor hallado en I, mediante sustitución hacia atrás (por ser U triangular superior).

En el caso general, y también por razones de estabilidad numérica, es necesario en general utilizar permutaciones para poder obtener una buena factorización LU, tal que $A = PLU$, con P una matriz de permutación.

Ejemplo 2.8.1 Sea

$$A = \begin{pmatrix} 2 & 4 & 2 \\ 1 & 5 & 2 \\ 4 & -1 & 9 \end{pmatrix}$$

Usando sólo operaciones por filas de Tipo III tenemos

$$\begin{pmatrix} 2 & 4 & 2 \\ 1 & 5 & 2 \\ 4 & -1 & 9 \end{pmatrix} \underset{\substack{f_2 - f_1/2 \\ f_3 - 2f_1}}{\longrightarrow} \begin{pmatrix} 2 & 4 & 2 \\ 0 & 3 & 1 \\ 0 & -9 & 5 \end{pmatrix} \underset{f_3 + 3f_2}{\longrightarrow} \begin{pmatrix} 2 & 4 & 2 \\ 0 & 3 & 1 \\ 0 & 0 & 8 \end{pmatrix} = U$$

De esta forma, $\alpha_1 = 1/2$, $\alpha_2 = 2$ y $\alpha_3 = -3$ por lo que utilizando (2.36), resulta

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 2 & -3 & 1 \end{pmatrix}$$

comprobándose que

$$LU = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 2 & -3 & 1 \end{pmatrix} \begin{pmatrix} 2 & 4 & 2 \\ 0 & 3 & 1 \\ 0 & 0 & 8 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 2 \\ 1 & 5 & 2 \\ 4 & -1 & 9 \end{pmatrix} = A$$

Esto es, la matriz A puede ser factorizada en un producto de una matriz triangular inferior L y otra matriz triangular superior U. Esto posee ventajas numéricas y es el método en el que se basan los programas usuales de resolución de sistemas lineales para matrices grandes.

En términos de matrices elementales, el proceso anterior puede ser representado como

$$E_3 E_2 E_1 A = U$$

donde

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{pmatrix}$$

comprobándose que

$$L = E_1^{-1} E_2^{-1} E_3^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 2 & -3 & 1 \end{pmatrix}$$

2.9. Algunas aplicaciones

2.9.1. Recuperación de información

Tarea: Buscar en una base de datos (una colección de miles o millones de documentos) - para encontrar alguno que se aproxime lo más posible a ciertos criterios de búsqueda.

Ejemplos de bases son: Páginas web, listas de archivos, libros, películas, etc.

- Supongamos que nuestra *base de datos* contiene m documentos, y
- se dispone de n palabras claves o frases para hacer la búsqueda (elegidas juiciosamente: evitando palabras simples y comunes o frases que no describan el contenido, como artículos, preposiciones, pronombres, etc.).

Entonces

- Ordenamos las palabras claves en forma alfabética (de 1 a n), y

- representamos la base de datos mediante una **matriz** A de $m \times n$, de la siguiente manera:
  (i) Las filas representan cada documento individualmente.
  (ii) Las columnas representan las palabras claves.
- a_{ij} = es la frecuencia relativa de encuentros de la j-ésima palabra clave en el i-ésimo documento.
- La lista de palabras claves que son usadas en una búsqueda específica se representan con un vector columna $\mathbf{x}$ en $\mathbb{R}^n$, donde

$$\begin{aligned} x_j &= 1 \quad \text{si la } j\text{-ésima palabra clave de la lista maestra} \\ & \text{está en nuestra búsqueda específica} \\ x_j &= 0 \quad \text{en caso contrario} \end{aligned}$$

- La búsqueda se realiza entonces al multiplicar A por el vector columna $\mathbf{x}$.

Ejemplo 2.9.1 Base de datos: libros de texto sobre Algebra Lineal.

1. Algebra lineal aplicada.
2. Algebra lineal elemental.
3. Algebra lineal elemental con aplicaciones.
4. Algebra lineal y sus aplicaciones.
5. Algebra lineal con aplicaciones.
6. Algebra de matrices con aplicaciones.
7. Teoría de matrices.

La colección de palabras claves es:

- *álgebra, aplicación, elemental, lineal, matriz, teoría*

Como los títulos de los libros no repiten ninguna palabra clave, podemos usar ceros y unos para los coeficientes a_{ij} de la matriz del ejemplo.

En general, las entradas a_{ij} podrán ser números enteros que representan la cantidad de veces que la palabra clave j aparece en el título o documento i).

- Asumimos que nuestra herramienta de búsqueda es suficientemente sofisticada y flexible como para identificar las diferentes formas de una misma palabra (aplicación = aplicaciones = aplicada).

- Los coeficientes para este caso serán

	Palabras claves					
Número del libro	álgebra	aplicación	elemental	lineal	matriz	teoría
(1)	1	1	0	1	0	0
(2)	1	0	1	1	0	0
(3)	1	1	1	1	0	0
(4)	1	1	0	1	0	0
(5)	1	1	0	1	0	0
(6)	1	1	0	0	1	0
(7)	0	0	0	0	1	1

Si nuestra búsqueda consiste en {aplicada, lineal, álgebra} entonces definimos el vector de búsqueda $\mathbf{x}$, y la matriz de la base de datos A:

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Ahora, buscamos $\mathbf{y} = \mathbf{Ax}$:

$$\mathbf{y} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 3 \\ 3 \\ 3 \\ 2 \\ 0 \end{pmatrix}$$

$$\begin{aligned} y_1 &= \text{cantidad de palabras-buscadas que coinciden en el título 1.} \\ y_2 &= \text{cantidad que coinciden en el título 2.} \\ &\vdots \\ y_m &= \text{cantidad que coinciden en el título } n. \end{aligned}$$

- Como $y_1 = y_3 = y_4 = y_5 = 3$, los libros $1, 3, 4, 5$ son los que mejor coinciden, porque contienen a las tres palabras claves buscadas.

  Si buscamos los títulos que contengan *todas* las palabras claves buscadas, entonces la respuesta es $1, 3, 4, 5$.

- En cambio, si buscamos los libros cuyos títulos contengan *al menos* una de las palabras claves buscadas, entonces la respuesta será: primero los 4 libros con 3 coincidencias, seguidos de los 2 libros con 2 coincidencias; en total 6 libros.

Una herramienta típica de búsqueda de alta performance puede buscar *millones* de documentos con *cientos de miles* de palabras claves posibles. No obstante, el problema de búsqueda es usualmente manejable ya que la matriz de la base de datos y los vectores de búsqueda son típicamente esparsos (contienen muchos ceros).

Las palabras claves de búsqueda deben ser elegidas con cuidado para optimizar el resultado: buscar en la Web libros de álgebra lineal usando las palabras claves *lineal* y *álgebra* podrá arrojar miles de aciertos, muchos de los cuales quizás no tengan nada que ver con álgebra lineal. A su vez, si usamos criterios muy restrictivos, podemos perder algunas páginas relevantes e interesantes.

Para páginas web, los coeficientes de la matriz de la base de datos deberían representar la frecuencia relativa de ocurrencias de la palabras claves en los documentos. Entonces, en vez de tratar de hacer coincidir todas las palabras de la lista de búsqueda extendida, podríamos dar prioridad a aquellas páginas/documentos que coincidan sobre todo en las de frecuencia relativa alta.

Para hacer esto necesitamos encontar las filas de la matriz A que estén más cerca del vector $\mathbf{x}$. Y para esto, necesitamos el concepto de ortogonalidad (que se tratará en detalle más adelante).

2.9.2. Redes y grafos

Uso de potencias de matrices en sistemas de comunicaciones

Tarea: Calcular la cantidad de caminos disponibles entre dos nodos de una red telefónica compleja.

La red telefónica se representa como un grafo: un conjunto de puntos $\{V_i\}$, llamados *vértices*, junto con un conjunto de pares (no ordenados) $\{V_i, V_j\}$, llamadas *aristas*. Esto es, un conjunto de puntos (por ej. los nodos de Internet), algunos de los cuales están conectados por líneas (por ej. fibra óptica).

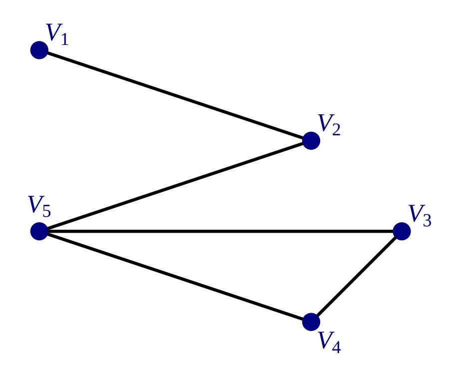


Figura 2.2: Ejemplo de grafo

Los segmentos de rectas que conectan los vértices corresponden a las aristas: $\{V_1, V_2\}$, $\{V_2, V_5\}$, $\{V_5, V_3\}$, $\{V_5, V_4\}$ y $\{V_3, V_4\}$.

Si tuviéramos miles (o millones) de aristas, el gráfico se podría complicar un poco.

Construimos la *matriz de representación* de una red: Si el grafo tiene un total de n vértices, se define una matriz $A = \{a_{ij}\}$ de $n \times n$:

$$a_{ij} = \begin{cases} 1 & \text{si existe la arista que une } V_i \text{ con } V_j \\ 0 & \text{si no existe una arista que una } V_i \text{ con } V_j. \end{cases}$$

La matriz A se llama la *matriz de adyacencia* o *matriz de vértices* del grafo.

En nuestro ejemplo sería

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

La matriz de adyacencia es simétrica ($a_{ij} = a_{ji}$) debido a que si V_i y V_j están conectados, entonces $a_{ij} = a_{ji} = 1$; y si no están conectados $a_{ij} = a_{ji} = 0$.

Consideremos un camino o senda en la grafo como una secuencia de aristas que unen un vértice con otro. En nuestro ejemplo, las aristas $\{V_1, V_2\}$ y $\{V_2, V_5\}$ representan un camino desde V_1 hasta V_5. El largo del camino o de la senda en este caso es 2 debido a que consiste de dos aristas.

Los camino se indican con flechas:

$$V_1 \longrightarrow V_2 \longrightarrow V_5$$

es un camino de longitud 2 desde V_1 hasta V_5. Y

$$V_4 \longrightarrow V_5 \longrightarrow V_2 \longrightarrow V_1$$

es un camino de longitud 3 desde V_4 hasta V_1.

Una arista puede atravesarse más de una vez en un mismo camino,

$$V_5 \longrightarrow V_3 \longrightarrow V_5 \longrightarrow V_3$$

es un camino de longitud 3 desde V_5 hasta V_3.

¿Cómo se puede usar la matriz de adyacencia para averiguar los caminos de diferentes longitudes (número de aristas que usan) que existen entre dos nodos particulares ?

Tomando potencias de la matriz de adyacencia podemos determinar el número de caminos (o sendas) de una longitud determinada entre dos vértices.

Esta información es crítica para lograr operaciones eficientes en sistemas de ruteo de telecomunicaciones de alta velocidad.

El siguiente teorema justifica la metodología.

Teorema 2.9.2: Sea A una matriz de adyacencia de $n \times n$ de un grafo. Si $a_{ij}^{(k)}$ representa el coeficiente en el lugar ij de su potencia A^k, entonces $a_{ij}^{(k)}$ es igual al número de caminos de longitud k entre los vértices V_i y V_j.

Demostración (por inducción). Para el caso $k = 1$, de la definición se sigue que los a_{ij} representan los caminos de longitud 1 entre V_i y V_j.

Supongamos ahora cierta la afirmación para un cierto valor m. Esto es, cada coeficiente de la matriz A^m representa el número de caminos de longitud m entre los vértices correspondientes ($a_{ij}^{(m)}$ es el número de caminos de longitud m entre V_i y V_j).

Si existe una arista $\{V_j, V_s\}$, entonces

$$a_{ij}^{(m)}.a_{js} = a_{ij}^{(m)}$$

es el número de caminos de longitud $(m+1)$ desde V_i hasta V_s de la forma

$$V_i \longrightarrow \cdots \longrightarrow V_j \longrightarrow V_s$$

Podemos calcular el total de caminos de longitud $(m+1)$ desde V_i hasta V_s de la siguiente manera:

$$a_{i1}^{(m)}.a_{1s} + a_{i2}^{(m)}.a_{2s} + \cdots + a_{in}^{(m)}.a_{ns}$$

Pero esta expresión representa efectivamente el coeficiente $a_{is}^{(m+1)}$ de la matriz $A^m.A = A^{m+1}$. ■

Ejemplo 2.9.2 Determine el número de caminos de longitud 3 entre cualesquiera dos vértices del grafo anterior.

$$A^3 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{pmatrix}^3 = \begin{pmatrix} 0 & 2 & 1 & 1 & 0 \\ 2 & 0 & 1 & 1 & 4 \\ 1 & 1 & 2 & 3 & 4 \\ 1 & 1 & 3 & 2 & 4 \\ 0 & 4 & 4 & 4 & 2 \end{pmatrix}$$

Por ejemplo, el número de caminos de longitud 3 entre los vértices V_3 y V_4 es $a_{34}^{(3)} = 3$.

Notar que A^3 también es simétrica: existe la misma cantidad de caminos de longitud 3 (o de cualquier longitud) desde V_i hasta V_j, que desde V_j hasta V_i.

Observar también los coeficientes de la diagonal principal y comparar con el grafo. Es imposible ir desde V_1 hasta V_1 (ni de V_2 a V_2) en 3 pasos. Por tanto, los correspondientes coeficientes de A^3 son nulos.

Otras aplicaciones de potencias de matrices se discutirán más adelante, cuando veampos el procedimiento óptimo para evaluar potencias arbitrarias de una matriz (basado en sus autovalores y autovectores).

Capítulo 3

Determinantes

3.1. Introducción

En el presente capítulo nos concentraremos en matrices A **cuadradas** $(n \times n)$, que son las que corresponden a sistemas lineales $A\mathbf{x} = \boldsymbol{b}$ con un número de ecuaciones igual al número de incógnitas. Hemos visto que tales matrices pueden ser de dos tipos:

I. A no singular. En este caso:

1. A es invertible ($\exists\ A^{-1}$)
2. El sistema $A\mathbf{x} = \boldsymbol{b}$ tiene solución **única** $\mathbf{x} = A^{-1}\boldsymbol{b}\ \forall\ \boldsymbol{b} \in \mathbb{R}^n$ (compatible determinado).
3. La forma escalonada reducida de Gauss-Jordan de la matriz es la matriz identidad de $n \times n$: $U = I_n$.

II. A singular. En este caso,

1. A no tiene inversa ($\nexists\ A^{-1}$)
2. El sistema $A\mathbf{x} = \boldsymbol{b}$ o bien tiene infinitas soluciones o bien no tiene solución (compatible indeterminado o incompatible)
3. La forma escalonada reducida de Gauss-Jordan de la matriz tiene al menos una fila nula.

Por lo tanto, frente a una matriz cuadrada A, la primer pregunta que surge es si es **singular o no singular**. Mostraremos aquí que existe un número obtenido a partir de los elementos de la matriz, llamado **determinante**, que discrimina estos dos casos: Es cero si la matriz es singular y distinto de cero si la matriz es no singular. Desde un punto de vista geométrico, el valor absoluto del determinante no es otra cosa que el "volumen" del "paralelepípedo" formado por las n filas o columnas de la matriz.

La idea de determinante es antigua, incluso anterior a la idea de matriz, ya que comenzó definiéndose como una propiedad del sistema de ecuaciones lineales. Al desarrollo del concepto y cálculo del determinante contribuyeron, entre otros, Gabriel Cramer, Alexandre Vandermonde, Pierre-Simon Laplace, Joseph-Louis Lagrange, Gottfried Leibniz, Carl F. Gauss, Augustin Cauchy, Carl G. Jacobi y James Sylvester.

El **objetivo** de este capítulo es entonces introducir un método simple para decidir si una matriz cuadrada A es singular o no-singular. El método permitirá, además, obtener una expresión analítica para la inversa de una matriz general de $n \times n$ no singular, y determinar en forma directa el "volumen" de un paralelepípedo general en 3 o más dimensiones.

3.2. Definición

Dada una matriz A de $n \times n$, deseamos definir un número $\det(A)$, función de los elementos a_{ij} de la matriz, que satisfaga

$$\begin{array}{rcl} \det(A) \neq 0 & \Rightarrow & A \text{ no singular} \\ \det(A) = 0 & \Rightarrow & A \text{ singular} \\ \det(I_n) & = & 1 \end{array} \tag{3.1}$$

De esta forma, $\det(A)$ "determinará" si la matriz es invertible o no invertible. La última condición (con I_n la matriz identidad de $n \times n$) fija la "escala" del determinante.

Una forma primaria de decidir si una matriz cuadrada A es singular o no es ver si su forma escalonada por filas U tiene ceros en la diagonal. Si los tiene es singular y si no los tiene es no singular (¡justificar!). Por lo tanto, si el producto de los elementos de la diagonal de U es cero, la matriz es singular, mientras que si el producto es distinto de cero, la matriz es no singular. El determinante de A estará entonces relacionado con este producto. Más aun, si A ya es triangular (inferior o superior), su determinante será, como veremos, directamente el producto de los elementos de su diagonal principal.

3.2.1. Casos básicos

- **Matrices de** 1×1. Si

  $$A = (a_{11})$$

  es una matriz de dimensión 1×1, entonces A tiene inversa si y sólo si $a_{11} \neq 0$, en cuyo caso $A^{-1} = \left(a_{11}^{-1}\right)$. Por lo tanto, para una matriz de 1×1 definimos

  $$\det(A) = a_{11}$$

  que verifica las tres condiciones (3.1). Por ejemplo, $\det(3) = 3$.

- **Matrices de** 2×2. Si

  $$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

  es de 2×2, hemos ya visto en los capítulos previos que A es no singular si y sólo si

  $$a_{11}a_{22} - a_{12}a_{21} \neq 0$$

  Repasemos el argumento: si $a_{11} \neq 0$, multiplicando la fila 2 por a_{11} y restando a este resultado la fila 1 multiplicada por a_{21}, se obtiene

  $$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \underset{a_{11}f_2 - a_{21}f_1}{\longrightarrow} \begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{11}a_{22} - a_{12}a_{21} \end{pmatrix}$$

  Esto muestra que si $a_{11} \neq 0$, A será equivalente por filas a I_2 (y por lo tanto tendrá inversa) si y sólo si $a_{11}a_{22} - a_{12}a_{21} \neq 0$.

Y si $a_{11}=0$, permutando las filas de A se obtiene

$$\begin{pmatrix} 0 & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \underset{f_1 \leftrightarrow f_2}{\longrightarrow} \begin{pmatrix} a_{21} & a_{22} \\ 0 & a_{12} \end{pmatrix}$$

por lo que en este caso A tendrá inversa si y sólo si $a_{21}a_{12} \neq 0$. Pero si $a_{11}=0$, esto es equivalente a $a_{11}a_{22} - a_{12}a_{21} \neq 0$.

Luego, si A es de 2×2 definimos

$$\det(A) = a_{11}a_{22} - a_{12}a_{21} \tag{3.2}$$

que verifica entonces las tres condiciones (3.1). La notación usualmente empleada es

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Ejemplo 3.2.1 Si $A = \begin{pmatrix} 2 & 4 \\ -3 & 1 \end{pmatrix}$,

$$\det(A) = \begin{vmatrix} 2 & 4 \\ -3 & 1 \end{vmatrix} = 2(1) - 4\,(-3) = 14$$

- **Matrices de** 3×3. Si

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

es una matriz de 3×3, podemos repetir el análisis de reducción por filas y mostrar que A es equivalente por filas a I_3 si y sólo si

$$a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \neq 0$$

Entonces definimos

$$\begin{aligned} \det(A) &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \end{aligned} \tag{3.3}$$

que verifica las tres condiciones (3.1).

La expresión (3.3) es la suma de $3! = 6$ **productos elementales** $a_{1j_1}a_{2j_2}a_{3j_3}$, donde (j_1, j_2, j_3) es un reordenamiento del conjunto $(1, 2, 3)$, con un signo $+1$ o -1 según sea el número de permutaciones respecto de $(1, 2, 3)$ par o impar.
Nótese que lo mismo ocurre para matrices de 2×2: $\det(A) = a_{11}a_{22} - a_{12}a_{21}$ contiene los $2! = 2$ productos elementales $a_{1j_1}a_{2j_2}$ (con $j_1 \neq j_2$) posibles en este caso, con el signo correspondiente. Y para $n = 1$, $\det(A) = a_{11}$ es el único término posible.

Nótese que si A es **triangular superior** ($a_{ij} = 0$ si $i > j$) o **inferior** ($a_{ij} = 0$ si $i < j$), las expresiones anteriores se reducen al **producto de los elementos de la diagonal:** $\det(A) = a_{11}a_{22}$ para A de 2×2 y $\det(A) = a_{11}a_{22}a_{33}$ para A de 3×3 (¡verificar!).

Problema 3.2.1 Probar que si $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix}$, entonces $\det(A) = (a_{11}a_{22} - a_{12}a_{21})a_{33}$, y que A es no singular si y sólo si $\det(A) \neq 0$.

3.2.2. Desarrollo por cofactores

Daremos ahora una expresión más sencilla y general para el determinante, que servirá para el caso $n \times n$.

- Considerando primero el caso 2×2, podemos escribir el determinante (3.2) como

$$\begin{aligned} \det(A) &= a_{11}a_{22} - a_{12}a_{21} \\ &= a_{11}\det(M_{11}) - a_{12}\det(M_{12}) \end{aligned} \tag{3.4}$$

  donde $M_{11} = (a_{22})$ es la submatriz de 1×1 obtenida al borrar la fila 1 y columna 1 de A, y $M_{12} = (a_{21})$ la submatriz obtenida al borrar la fila 1 y columna 2 de A:

$$\begin{pmatrix} \cancel{a_{11}} & \cancel{a_{12}} \\ \cancel{a_{21}} & a_{22} \end{pmatrix} \longrightarrow M_{11} = \begin{pmatrix} a_{22} \end{pmatrix}, \quad \begin{pmatrix} \cancel{a_{11}} & \cancel{a_{12}} \\ a_{21} & \cancel{a_{22}} \end{pmatrix} \longrightarrow M_{12} = \begin{pmatrix} a_{21} \end{pmatrix}$$

- Para el caso de 3×3, se puede reordenar la expresión (3.3) y escribirla como

$$\begin{aligned} \det(A) &= a_{11}\,(a_{22}a_{33} - a_{23}a_{32}) - a_{12}\,(a_{21}a_{33} - a_{23}a_{31}) + a_{13}\,(a_{21}a_{32} - a_{22}a_{31}) \\ &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11}\det(M_{11}) - a_{12}\det(M_{12}) + a_{13}\det(M_{13}) \end{aligned} \tag{3.5}$$

  donde M_{1j} es la matriz de 2×2 obtenida al borrar la fila 1 y la columna j de A:

$$\begin{pmatrix} \cancel{a_{11}} & \cancel{a_{12}} & \cancel{a_{13}} \\ \cancel{a_{21}} & a_{22} & a_{23} \\ \cancel{a_{31}} & a_{32} & a_{33} \end{pmatrix} \longrightarrow M_{11} = \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}$$

$$\begin{pmatrix} \cancel{a_{11}} & \cancel{a_{12}} & \cancel{a_{13}} \\ a_{21} & \cancel{a_{22}} & a_{23} \\ a_{31} & \cancel{a_{32}} & a_{33} \end{pmatrix} \longrightarrow M_{12} = \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix}$$

$$\begin{pmatrix} \cancel{a_{11}} & \cancel{a_{12}} & \cancel{a_{13}} \\ a_{21} & a_{22} & \cancel{a_{23}} \\ a_{31} & a_{32} & \cancel{a_{33}} \end{pmatrix} \longrightarrow M_{13} = \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

Ejemplo 3.2.2. Si $A = \begin{pmatrix} 2 & 5 & 4 \\ 3 & 1 & 2 \\ 5 & 4 & 6 \end{pmatrix}$, entonces

$$\begin{aligned} \det(A) = \begin{vmatrix} 2 & 5 & 4 \\ 3 & 1 & 2 \\ 5 & 4 & 6 \end{vmatrix} &= 2\begin{vmatrix} 1 & 2 \\ 4 & 6 \end{vmatrix} - 5\begin{vmatrix} 3 & 2 \\ 5 & 6 \end{vmatrix} + 4\begin{vmatrix} 3 & 1 \\ 5 & 4 \end{vmatrix} \\ &= 2\,(6-8) - 5\,(18-10) + 4\,(12-5) \\ &= -16 \end{aligned}$$

Definición.
Dada A de $n \times n$, sea M_{ij} la submatriz de $(n-1) \times (n-1)$ obtenida al borrar la fila i y la columna j de A. Entonces

(i) El número $\det(M_{ij})$ se denomina **menor** del elemento a_{ij}.

(ii) El número $c_{ij} = (-1)^{i+j} \det(M_{ij})$ se denomina **cofactor** de a_{ij}.

- Para matrices de 2×2, podemos ahora reescribir (3.4) como

$$\det(A) = a_{11}a_{22} - a_{12}a_{21} = a_{11}c_{11} + a_{12}c_{12}$$

Esta es la **expansión por cofactores** de $\det(A)$ a lo largo de la **fila** 1.
Pero podemos también escribir el mismo determinante como

$$\det(A) = a_{21}\,(-a_{12}) + a_{22}a_{11} = a_{21}c_{21} + a_{22}c_{22}$$

Esta es la expansión por cofactores de $\det(A)$ a lo largo de la **fila** 2.

Usando ahora las columnas en lugar de las filas, podemos también realizar una expansión por cofactores del mismo $\det(A)$ a lo largo de cualquiera de las columnas:

$$\begin{aligned} \det(A) &= a_{11}a_{22} + a_{21}\,(-a_{12}) = a_{11}c_{11} + a_{21}c_{21} \qquad \text{(primer columna)} \\ &= a_{12}\,(-a_{21}) + a_{22}a_{11} = a_{12}c_{12} + a_{22}c_{22} \qquad \text{(segunda columna)} \end{aligned}$$

- Similarmente, para A de 3×3, la expresión (3.5) puede escribirse como

$$\det(A) = a_{11}c_{11} + a_{12}c_{12} + a_{13}c_{13} \quad \text{(expansión por fila 1)}$$

Y al igual que en el caso 2×2, el mismo determinante anterior se puede expandir a lo largo de *cualquier fila* o *cualquier columna*:

$$\begin{aligned} \det(A) &= a_{i1}c_{i1} + a_{i2}c_{i2} + a_{i3}c_{i3} \quad \text{(expansión por fila } i,\ i = 1,2,3) \\ \det(A) &= a_{1j}c_{1j} + a_{2j}c_{2j} + a_{3j}c_{3j} \quad \text{(expansión por columna } j,\ j = 1,2,3) \end{aligned}$$

Ejemplo 3.2.3. Calculemos el determinante de la matriz A del ejemplo anterior a lo largo de la columna 2:

$$\begin{aligned} \det(A) = \begin{vmatrix} 2 & 5 & 4 \\ 3 & 1 & 2 \\ 5 & 4 & 6 \end{vmatrix} &= 5(-1)^3 \begin{vmatrix} 3 & 2 \\ 5 & 6 \end{vmatrix} + 1(-1)^4 \begin{vmatrix} 2 & 4 \\ 5 & 6 \end{vmatrix} + 4(-1)^5 \begin{vmatrix} 2 & 4 \\ 3 & 2 \end{vmatrix} \\ &= -5\,(18-10) + (12-20) - 4\,(4-12) \\ &= -16 \end{aligned}$$

3.2.3. El caso general $n \times n$

Definición inductiva

Partiendo del caso trivial 1×1,

$$\det(A) = a_{11} \quad (n = 1)$$

para A de $n \times n$, $n > 1$, podemos definir el determinante como

$$\det(A) = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = a_{11}c_{11} + a_{12}c_{12} + \dots + a_{1n}c_{1n} \quad (n > 1) \tag{3.6}$$

donde

$$c_{1j} = (-1)^{1+j} \det(M_{1j}) \,, \quad j = 1, \dots, n$$

son los cofactores de los coeficientes de la primer fila de A.

Se puede utilizar también **cualquier fila o columna** de A para las expansiones:

$$\begin{aligned} \det(A) &= a_{i1}c_{i1} + a_{i2}c_{i2} + \dots + a_{in}c_{in} \quad \text{expansión por fila } i & (3.7) \\ &= a_{1j}c_{1j} + a_{2j}c_{2j} + \dots + a_{nj}c_{nj} \quad \text{expansión por columna } j & (3.8) \end{aligned}$$

donde

$$c_{ij} = (-1)^{i+j} \det(M_{ij})$$

es el cofactor del elemento a_{ij} de A.

De esta forma el determinante de una matriz de $n \times n$ queda expresado como una combinación de n determinantes de orden $(n - 1)$.

En la práctica, los determinantes se expanden a lo largo de la fila o columna que contenga más ceros (¿por qué?). Pueden utilizarse también otras propiedades del determinante para su evaluación, como veremos a continuación.

Forma explícita
Puede probarse que la definición (3.7) conduce a una función $\det(A) : \mathbb{R}^{n\times n} \to \mathbb{R}$, que es la suma de todos los $n!$ productos elementales de n elementos $a_{1j_1}a_{2j_2}\ldots a_{nj_n}$ (uno por cada fila y columna) de A, con un signo $+$ o $-$ de acuerdo al número de permutaciones $N_{j_1\ldots j_n}$ necesarias para llevar $(j_1, j_2, \ldots, j_n)$ al orden normal $(1, 2, \ldots, n)$:

$$\det(A) = \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} = \sum_{\substack{j_1,j_2,\ldots,j_n \\ j_i\neq j_k \text{ si } i\neq k}} (-1)^{N_{j_1\ldots j_n}} a_{1j_1}\ldots a_{nj_n} \qquad (3.9)$$
$$= a_{11}a_{22}\ldots a_{n-1,n-1}a_{nn} - a_{11}a_{22}\ldots a_{n-1,n}a_{n,n-1} + \ldots$$

La expresión (3.9) generaliza al caso $n \times n$ la fórmula explícita (3.3) para matrices de 3×3.
Si A es triangular superior o inferior, (3.9) se reduce a $\det(A) = a_{11}\ldots a_{nn}$. (¡probar!)

Ejemplo 3.2.4 Sea $A = \begin{pmatrix} 0 & 2 & 3 & 0 \\ 0 & 4 & 5 & 0 \\ 0 & 1 & 0 & 3 \\ 2 & 0 & 1 & 3 \end{pmatrix}$.
Conviene utlizar la primer columna para la expansión por cofactores, y luego la tercer columna para evaluar el menor $\det(M_{41})$:

$$\det(A) = \begin{vmatrix} 0 & 2 & 3 & 0 \\ 0 & 4 & 5 & 0 \\ 0 & 1 & 0 & 3 \\ 2 & 0 & 1 & 3 \end{vmatrix} = -2\begin{vmatrix} 2 & 3 & 0 \\ 4 & 5 & 0 \\ 1 & 0 & 3 \end{vmatrix} = -2\,.3\begin{vmatrix} 2 & 3 \\ 4 & 5 \end{vmatrix} = -2\,.3\,(10-12) = 12$$

Problema 3.2.2. Evaluar los determinantes de las siguientes matrices:

$$\text{a) } \begin{pmatrix} 2 & 1 \\ -1 & 3 \end{pmatrix}, \quad \text{b) } \begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \text{c) } \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

3.3. Propiedades del determinante

1. El determinante de una matriz A triangular (superior o inferior) es el producto de los n elementos de su diagonal principal:

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & 0 & \ldots & 0 \\ a_{21} & a_{22} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \ldots & \ldots & a_{nn} \end{vmatrix} = a_{11}a_{22}\ldots a_{nn}$$

Esto incluye en particular el caso en que A es diagonal ($a_{ij} = 0$ si $a \neq j$).

Este resultado ya lo hemos mostrado en base a la forma explícita. A partir de la definición recursiva, el resultado es obvio a partir del desarrollo de $\det(A)$ por la

columna 1 (A triangular superior) o fila 1 (A triangular inferior):

$$\det(A) = a_{11}\det M_{11} = a_{11}a_{22}\det M'_{11} = \ldots = a_{11}a_{22}\ldots a_{nn}$$

Por ejemplo,

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 6 & 7 \\ 0 & 0 & 8 & 9 \\ 0 & 0 & 0 & 10 \end{vmatrix} = 1.\begin{vmatrix} 5 & 6 & 7 \\ 0 & 8 & 9 \\ 0 & 0 & 10 \end{vmatrix} = 1\,.5\begin{vmatrix} 8 & 9 \\ 0 & 10 \end{vmatrix} = 1\,.5\,.8\,.10 = 400$$

2. El determinante de la traspuesta A^T es igual al de A:

$$\det\left(A^T\right) = \det(A)$$

Para $n = 1$ es obviamente válida. Asumiendo luego que la propiedad es válida para matrices de $(n-1)\times(n-1)$, y recordando que las filas de A^T son las columnas de A, vemos que el desarrollo por la fila i de $\det(A^T)$ coincide con el desarrollo por la columna i de $\det(A)$, siendo por lo tanto iguales. Por ejemplo,

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 3 \\ 2 & 4 \end{vmatrix} = 4 - 6 = -2$$

3. Si A tiene una fila o una columna nula entonces $\det(A) = 0$.
Es inmediato, considerando que el determinante se puede desarrollar por esa fila o columna. Por ejemplo, $\begin{vmatrix} 0 & 0 \\ 1 & 2 \end{vmatrix} = 0$.

4. Si B se obtiene de A intercambiando dos filas (o columnas) $\Rightarrow \det(B) = -\det(A)$.
Supongamos que se intercambia la fila i con la $i+1$. Entonces el desarrollo por fila $i+1$ de $\det(B)$ coincidirá con el desarrollo por fila i de $\det(A)$, excepto por un cambio de signo ($(-1)^{i+1+j} = -(-1)^{i+j}$). Para intercambios $i \leftrightarrow k$ con $k \neq i$, el resultado puede verse en forma análoga, o también considerando sucesivos intercambios $i \leftrightarrow i+1$ con la fila contigua: si $k > i$, se necesitan $k - i$ intercambios de este tipo para llevar la fila i a la k, y luego $k-1-i$ intercambios para llevar la ex-fila k (que quedó en la posición $k-1$) a la i. El signo total es $(-1)^{k-i+k-1-i} = -1$. Si se intercambian dos columnas la demostración es similar (o puede verse por su traspuesta). Por ejemplo,

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = -\begin{vmatrix} 3 & 4 \\ 1 & 2 \end{vmatrix} = -\begin{vmatrix} 2 & 1 \\ 4 & 3 \end{vmatrix} = -2$$

5. Si A tiene dos filas o dos columnas idénticas entonces $\det(A) = 0$.
Se puede obtener fácilmente como consecuencia de la propiedad anterior. Basta con intercambiar esas dos filas o columnas: se obtendrá $\det(A) = -\det(A)$, por lo que $\det(A) = 0$. Por ejemplo, $\begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix} = -\begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix}$, por lo que $\begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix} = 0$.

6. Si B se obtiene de A multiplicando una fila (o columna) de A por $\alpha \Rightarrow$ $\det(B) = \alpha \det(A)$:

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ \alpha a_{i1} & \dots & \alpha a_{in} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \alpha \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}$$

Esto es inmediato a partir del desarrollo de $\det(B)$ por esa fila (o columna): $\det(B) = \sum_j (\alpha a_{ij})(-1)^{i+j} \det M_{ij} = \alpha \sum_j a_{ij}(-1)^{i+j} \det M_{ij} = \alpha \det(A)$. Por ej.,

$$\begin{vmatrix} 2 & 4 \\ 3 & 1 \end{vmatrix} = 2 \begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} = -10$$

7. $\det(\alpha A) = \alpha^n \det(A)$ si A es de $n \times n$.
Este resultado se obtiene aplicando el anterior n veces, es decir, a todas las filas (o columnas). Por ejemplo,

$$\begin{vmatrix} 2 & 4 \\ 6 & 2 \end{vmatrix} = 2^2 \begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} = -20$$

8. Si B se obtiene de A sumando a una fila (o columna) de A un múltiplo de otra fila (o columna) de $A \Rightarrow \det(B) = \det(A)$:

$$\begin{vmatrix} a_{11} + \alpha a_{i1} & \dots & a_{1n} + \alpha a_{in} \\ a_{21} & \dots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}$$

Se demuestra mediante el desarrollo de $\det(B)$ por la fila modificada. Se obtendrá $\det(B) = \det(A) + \alpha \det(\text{matriz con dos filas iguales}) = \det(A)$. Por ejemplo,

$$\begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix} \underset{f_2 - 2f_1}{\longrightarrow} \begin{pmatrix} 1 & 3 \\ 0 & -1 \end{pmatrix} \Rightarrow \begin{vmatrix} 1 & 3 \\ 2 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 3 \\ 0 & -1 \end{vmatrix} = -1$$

9. Si A tiene dos filas o columnas proporcionales $\Rightarrow \det(A) = 0$.
En efecto, si la fila i es α veces la fila j, con $j \neq i$, el desarrollo por la fila i implica $\det(A) = \alpha \det(\text{matriz con dos filas iguales}) = 0$.
Generalizando, $\det(A) = 0$ si una de las filas (o columnas) es combinación lineal de otras filas (columnas). Esto significa que dicha fila es una suma de otras filas de A multiplicadas por constantes. El desarrollo de $\det(A)$ por dicha fila será una combinación lineal de determinantes de matrices con dos filas iguales, por lo que será 0. Por ejemplo, si la fila 3 es la fila 1 m´as dos veces la fila 2, el desarrollo por la fila 3 conduce a (¡verificar!)

$$\begin{vmatrix} a & b & c \\ d & e & f \\ a + 2d & b + 2e & c + 2f \end{vmatrix} = \begin{vmatrix} a & b & c \\ d & e & f \\ a & b & c \end{vmatrix} + 2 \begin{vmatrix} a & b & c \\ d & e & f \\ d & e & f \end{vmatrix} = 0 + 0 = 0 \qquad (3.10)$$

10. Notar que en general, $\det(A+B) \neq \det(A) + \det(B)$.
Por ejemplo, si

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad \Rightarrow \quad A+B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

y se obtiene $\det(A) = \det(B) = 0$ mientras que $\det(A+B) = \det(I_2) = 1$.

Ejemplo 3.3.1 Aplicando estas propiedades, podemos calcular determinantes sin utilizar explícitamente el desarrollo por cofactores, llevando la matriz a la forma triangular. Por ejemplo, si

$$A = \begin{pmatrix} 0 & 1 & 5 \\ 3 & -6 & 9 \\ 2 & 6 & 1 \end{pmatrix}$$

paso 1	$\begin{vmatrix} 0 & 1 & 5 \\ 3 & -6 & 9 \\ 2 & 6 & 1 \end{vmatrix} = -\begin{vmatrix} 3 & -6 & 9 \\ 0 & 1 & 5 \\ 2 & 6 & 1 \end{vmatrix}$	intercambio de fila 1 y 2 (prop. 4)
paso 2	$= -3\begin{vmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 2 & 6 & 1 \end{vmatrix}$	se extrae un factor 3 de la fila 1 (prop. 6)
paso 3	$= -3\begin{vmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 0 & 10 & -5 \end{vmatrix}$	se suma -2 veces la fila 1 a la fila 3 (prop. 8)
paso 4	$= -3\begin{vmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & -55 \end{vmatrix}$	se suma -10 veces la fila 2 a la fila 3 (prop. 8)
paso 5	$= (-3)(1\,.1\,.(-55))$	determinante de matriz triangular (prop. 1)
	$= 165$	

Ejemplo 3.3.2 Si $A = \begin{pmatrix} 1 & 3 & -2 & 4 \\ 2 & 6 & -4 & 8 \\ 3 & 9 & 1 & 5 \\ 1 & 1 & 4 & 8 \end{pmatrix}$,

$$\det(A) = \begin{vmatrix} 1 & 3 & -2 & 4 \\ 2 & 6 & -4 & 8 \\ 3 & 9 & 1 & 5 \\ 1 & 1 & 4 & 8 \end{vmatrix} = 2\begin{vmatrix} 1 & 3 & -2 & 4 \\ 1 & 3 & -2 & 4 \\ 3 & 9 & 1 & 5 \\ 1 & 1 & 4 & 8 \end{vmatrix} = 0$$

ya que la última matriz tiene dos filas iguales (prop. 5). Se llega al mismo resultado aplicando directamente la propiedad 9 (fila 2 $\propto$ fila 1) o restando a la fila 2 dos veces la fila 1, que anula la fila 2.

Problemas 3.3

1. Evaluar los determinantes de las siguientes matrices utilizando las propiedades anteriores (llevarlas, por ejemplo, a una forma triangular). Indique cuales son singulares.

$$\text{a)}\ \begin{pmatrix}-1 & 3 & 2\\ 2 & 1 & 3\\ 1 & 4 & 5\end{pmatrix},\quad \text{b)}\ \begin{pmatrix}1 & 2 & 1\\ 1 & 2 & 0\\ 1 & 0 & 0\end{pmatrix},\quad \text{c)}\ \begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 9\end{pmatrix}\quad \text{d)}\ \begin{pmatrix}1 & 1 & 1 & 1\\ -1 & 1 & 1 & 1\\ 1 & -1 & 1 & 1\\ 1 & 1 & -1 & 1\end{pmatrix}$$

2. Probar que
$$\begin{vmatrix}1 & 1 & 1\\ a & b & c\\ a^2 & b^2 & c^2\end{vmatrix} = (b-a)(c-a)(c-b)$$
(Es el caso 3×3 del determinante de Vandermonde, que aparece en varias aplicaciones).

3. Probar que la ecuación de una recta (en el plano) que pasa por (x_1, y_1) y (x_2, y_2) puede expresarse como
$$\begin{vmatrix}x & y & 1\\ x_1 & y_1 & 1\\ x_2 & y_2 & 1\end{vmatrix} = 0$$

3.4. Aplicaciones geométricas del determinante

1. Interpretación geométrica del determinante. Caso 2×2.
Dos vectores $\boldsymbol{u} = (a, b)$, $\boldsymbol{v} = (c, d)$ en el plano forman los lados de un paralelogramo (ver figura 3.1). Dado que
$$a = |\boldsymbol{u}|\cos\alpha,\quad b = |\boldsymbol{u}|\operatorname{sen}\alpha,\quad c = |\boldsymbol{v}|\cos\beta,\quad d = |\boldsymbol{v}|\operatorname{sen}\beta$$
vemos que el determinante es
$$\begin{aligned}\begin{vmatrix}a & b\\ c & d\end{vmatrix} = ad - bc &= |\boldsymbol{u}||\boldsymbol{v}|(\cos\alpha\operatorname{sen}\beta - \operatorname{sen}\alpha\cos\beta)\\ &= |\boldsymbol{u}||\boldsymbol{v}|\sin(\beta-\alpha) = |\boldsymbol{u}|h \end{aligned}\qquad (3.11)$$
siendo $h = |\boldsymbol{v}|\sin(\beta-\alpha)$ la “altura” del paralelogramo. Pero $|\boldsymbol{u}|h$ es justamente el área del paralelogramo. Por lo tanto (y considerando que segun el orden elegido, $\beta-\alpha$ puede ser ≥ 0 o ≤ 0) tenemos en general
$$\text{Area} = |\boldsymbol{u}||\boldsymbol{v}||\sin(\beta-\alpha)| = |ad-bc| = |\det(A)| \qquad (3.12)$$
con $A = \begin{pmatrix}a & b\\ c & d\end{pmatrix}$. Si $\beta = \alpha$ los vectores son colineales ($\boldsymbol{v}\propto\boldsymbol{u}$) y Area$=\det(A) = 0$.

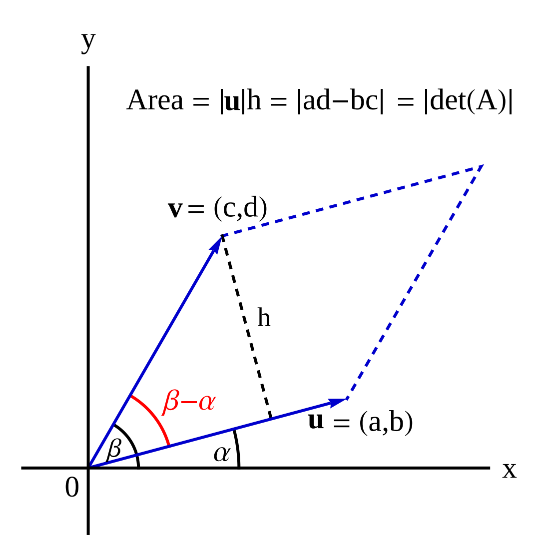


Figura 3.1: Determinante de una matriz de 2×2. Su valor absoluto representa el área del paralegramo formado por sus filas (o columnas).

2. Producto vectorial. Dados dos vectores de $\mathbb{R}^3$,

$$\begin{array}{rcl} \boldsymbol{u} &=& (u_1,u_2,u_3) &=& u_1\boldsymbol{e}_1+u_2\boldsymbol{e}_2+u_3\boldsymbol{e}_3\\ \boldsymbol{v} &=& (v_1,v_2,v_3) &=& v_1\boldsymbol{e}_1+v_2\boldsymbol{e}_2+v_3\boldsymbol{e}_3\end{array}$$

donde $\boldsymbol{e}_1=(1,0,0)$, $\boldsymbol{e}_2=(0,1,0)$, $\boldsymbol{e}_3=(0,0,1)$, el producto vectorial (o cruz) $\boldsymbol{u}\times\boldsymbol{v}$ se define como el vector

$$\boldsymbol{u}\times\boldsymbol{v} = \begin{vmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3\\ u_1 & u_2 & u_3\\ v_1 & v_2 & v_3\end{vmatrix} \tag{3.13}$$

$$= (u_2v_3-u_3v_2)\boldsymbol{e}_1+(u_3v_1-u_1v_3)\boldsymbol{e}_2+(u_1v_2-u_2v_1)\boldsymbol{e}_3 \tag{3.14}$$

Se deja como problema probar que $\boldsymbol{u}\times\boldsymbol{v}$ es **ortogonal** (perpendicular) a $\boldsymbol{u}$ y a $\boldsymbol{v}$ (véase primero el punto 3. siguiente) y que

$$|\boldsymbol{u}\times\boldsymbol{v}| = |\boldsymbol{u}||\boldsymbol{v}||\sin\theta| \tag{3.15}$$

siendo θ el ángulo entre $\boldsymbol{u}$ y $\boldsymbol{v}$.
(muestre (3.15) primero para vectores $\boldsymbol{u}$ y $\boldsymbol{v}$ contenidos en el plano x,y; puede luego extender el resultado probando que en general, $|\boldsymbol{u}\times\boldsymbol{v}|^2=|\boldsymbol{u}|^2|\boldsymbol{v}|^2-(\boldsymbol{u}\cdot\boldsymbol{v})^2$).

El resultado (3.15) muestra también que $|\boldsymbol{u}\times\boldsymbol{v}|$ es el **área** del paralelogramo formado por los vectores $\boldsymbol{u}$ y $\boldsymbol{v}$ (véase (3.12)).

3. Producto triple.
Si $\boldsymbol{w}=(w_1,w_2,w_3)$, el producto escalar $\boldsymbol{w}\cdot(\boldsymbol{u}\times\boldsymbol{v})$ (producto triple) puede expresarse

como un determinante:

$$\begin{aligned}\boldsymbol{w}\cdot(\boldsymbol{u}\times\boldsymbol{v}) &= w_1(\boldsymbol{u}\times\boldsymbol{v})_1+w_2(\boldsymbol{u}\times\boldsymbol{v})_2+w_3(\boldsymbol{u}\times\boldsymbol{v})_3\\ &= \begin{vmatrix} w_1 & w_2 & w_3\\ u_1 & u_2 & u_3\\ v_1 & v_2 & v_3\end{vmatrix}\end{aligned} \tag{3.16}$$

Este resultado es obvio a partir de (3.13)–(3.14). Se dejan los detalles para el lector. Nótese que si $\boldsymbol{w}=\boldsymbol{u}$ o $\boldsymbol{w}=\boldsymbol{v}\Rightarrow$ (3.16) es nulo.

4. Interpretación geométrica del determinante. Caso 3×3.
El producto triple anterior puede también escribirse como

$$\boldsymbol{w}\cdot(\boldsymbol{u}\times\boldsymbol{v}) = |\boldsymbol{w}||\boldsymbol{u}\times\boldsymbol{v}|\cos\phi = h|\boldsymbol{u}\times\boldsymbol{v}| \tag{3.17}$$

donde ϕ es el ángulo entre $\boldsymbol{w}$ y $\boldsymbol{u}\times\boldsymbol{v}$ y $h=|\boldsymbol{w}|\cos\phi$ la "altura" del paralelepípedo formado por los tres vectores $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ (ver figura, en la que $\boldsymbol{u}$ y $\boldsymbol{v}$ están en el plano x,y). Como $|\boldsymbol{u}\times\boldsymbol{v}|$ es el área de la base, el módulo del producto anterior es el **volumen** del paralelepípedo:

$$\text{Volumen} = |\boldsymbol{w}\cdot(\boldsymbol{u}\times\boldsymbol{v})| = |\det(A)|,\quad A=\begin{pmatrix} w_1 & w_2 & w_3\\ u_1 & u_2 & u_3\\ v_1 & v_2 & v_3\end{pmatrix} \tag{3.18}$$

Los vectores $\boldsymbol{w}$, $\boldsymbol{u}$, $\boldsymbol{v}$ pueden ponerse también por columna ya que $\det(A)=\det(A^T)$.

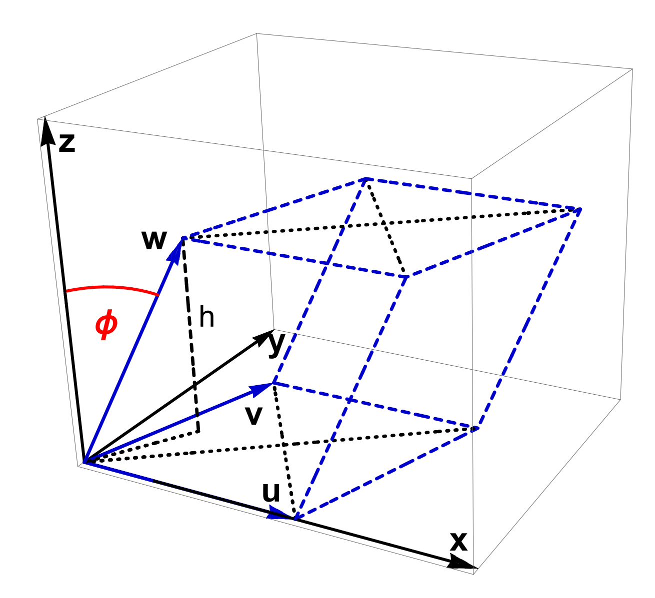


Figura 3.2: Determinante de una matriz de 3×3. Su valor absoluto representa el volumen del paralelepípedo formado por sus filas (o columnas).

La noción de volumen pueden generalizarse a $\mathbb{R}^n$, siendo $|\det(A)|$ el "volumen" (o hipervolumen) del "paralelepípedo" formado por las n filas o columnas de A.

5. Jacobiano.
   Al hacer un cambio de variables en integrales dobles, triples, etc., el Jacobiano de la transformación se calcula mediante un determinante. Así, en $\mathbb{R}^3$, si $x = g(u, v, t)$, $y = h(u, v, t)$ y $z = l(u, v, t)$, el Jacobiano de la transformación es,

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v} \\ \frac{\partial x}{\partial t} & \frac{\partial y}{\partial t} & \frac{\partial z}{\partial t} \end{vmatrix}$$

   Verificar que si $x = r\cos\theta$, $y = r\,\text{sen}\,\theta$ y $z = z$ (coordenadas cilíndricas) $J = r$ mientras que si $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, $z = r\cos\theta$ (coordenadas esféricas) $J = r^2\sin\theta$. Interprete geométricamente estos resultados.

Problemas 3.4.1

1. Determinar el área del paralelogramo determinado por los vectores
   a) $\{(1, 0), (1, 1)\}$, b) $\{(a, 0), (b, c)\}$
   En b), explicar porqué el área no depende de b.

2. Determinar el volumen del paralelepípedo formado por los vectores
   a) $\{(1, 0, 0), (1, 1, 0), (1, 1, 1)\}$, b) $\{(a, 0, 0), (b, c, 0), (d, e, f)\}$
   c) $\{(1, 1, 1), (-1, 1, 1), (1, -1, 1)\}$.
   En b) explicar porqué el volumen no depende de b, d, e.

3. Muestre que el volumen generado por las filas y el generado por las columnas de una matriz A (de 3×3) son iguales. Muestre también que el mismo es nulo si y sólo si A es singular.

4. Probar que $|\boldsymbol{u} \times \boldsymbol{v}|$ es el área del paralelogramo formado por $\boldsymbol{u}$ y $\boldsymbol{v}$.

5. A partir de las propiedades del determinante, probar que
   a) $\boldsymbol{u} \times \boldsymbol{v}$ es perpendicular a $\boldsymbol{u}$ y $\boldsymbol{v}$.
   b) $\boldsymbol{w} \cdot (\boldsymbol{u} \times \boldsymbol{v}) = \boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w}) = \boldsymbol{v} \cdot (\boldsymbol{w} \times \boldsymbol{u})$.

6. Muestre que el producto $\boldsymbol{u} \times \boldsymbol{v}$ no es conmutativo ni asociativo.
7. Determine el Jacobiano para coordenadas cilíndricas y esféricas.

3.5. Resultados claves

3.5.1. Determinante de matrices elementales

Las propiedades 4., 6. y 8. de la sección 3.3 pueden expresarse en términos de las matrices elementales correspondientes a las operaciones de Tipo I, II y III que se vieron en el método de eliminación gaussiana.

Operación de Tipo I: Intercambiar dos filas de A ($A \longrightarrow E_I\, A$), donde E_I es, por ejemplo,

$$E_I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Se verifica que

$$\begin{aligned} \det(E_I A) &= -\det(A) \\ &= \det(E_I)\det(A) \quad (\text{ya que } \det(E_I) = -\det(I) = -1) \end{aligned}$$

Operación de Tipo II: Multiplicar una fila por un escalar $\alpha \neq 0$ ($A \longrightarrow E_{II}\, A$), donde E_{II} es, por ejemplo,

$$E_{II} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Se verifica que

$$\begin{aligned} \det(E_{II} A) &= \alpha \det(A) \\ &= \det(E_{II})\det(A) \quad (\text{ya que } \det(E_{II}) = \alpha . \det(I) = \alpha) \end{aligned}$$

Operación de Tipo III: Sumar a una fila, un múltiplo de otra fila ($A \longrightarrow E_{III}\, A$), donde E_{III} es, por ejemplo,

$$E_{III} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \alpha \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Aplicando la propiedad 8. a E_{III} y a A

$$\begin{aligned} \det(E_{III} A) &= \det(A) \\ &= \det(E_{III})\det(A) \quad (\text{ya que } \det(E_{III}) = \det(I) = 1) \end{aligned}$$

Resumiendo, si E es una matriz elemental, entonces

$$\det(EA) = \det(E)\det(A)$$

siendo

$$\det(E) = \begin{cases} -1 & \text{si } E \text{ es de Tipo I} \\ \alpha & \text{si } E \text{ es de Tipo II} \\ 1 & \text{si } E \text{ es de Tipo III} \end{cases}$$

Lo mismo sucede para columnas.
Recordar que las operaciones elementales por columnas se obtienen al multiplicar a la derecha por una matriz elemental E_I, E_{II} o E_{III}. Usando la propiedad 2.,

$$\begin{aligned} \det(AE) &= \det\left((AE)^T\right) \\ &= \det\left(E^T A^T\right) \\ &= \det\left(E^T\right)\det\left(A^T\right) \\ &= \det(E)\det(A) \end{aligned}$$

Los efectos sobre el determinante debido a las operaciones sobre las columnas son idénticos a los correspondientes a las operaciones sobre las filas.

3.5.2. Determinante de matrices singulares y de un producto

Demostraremos ahora dos propiedades fundamentales del determinante.

¡Importante!

1. Una matriz A de dimensión $n \times n$ es singular si y sólo si $\det(A) = 0$.
2. Si A y B son matrices de $n \times n$,

$$\det(AB) = \det(A)\det(B)$$

Primer resultado clave:

1. Una matriz A de dimensión $n \times n$ es singular si y sólo si $\det(A) = 0$.

Demostración. Cualquier matriz A de $n \times n$ se puede reducir, mediante una cantidad finita de operaciones elementales sobre las filas, a la forma escalonada reducida de Gauss-Jordan U (si A es no singular, $U = I_n$, mientras que si es A es singular, U tiene al menos una fila nula):

$$U = E_k E_{k-1} \dots E_1 A$$

donde todas las $\{E_i\}$ son matrices elementales. Además, por ser productos con matrices elementales se sabe que

$$\begin{aligned} \det(U) &= \det(E_k E_{k-1} ... E_1 A) \\ &= \det(E_k)\det(E_{k-1})\dots\det(E_1)\det(A) \end{aligned} \qquad (3.19)$$

Pero $\det(E_i)$ es siempre no nulo (Tipo I, II o III), luego

$$\det(A) = 0 \text{ si y sólo si } \det(U) = 0$$

Cuando A es singular (A no equivalente por filas a I_n) U tiene al menos una fila nula y por lo tanto

$$\det(U) = 0$$

que es equivalente a $\det(A) = 0$.

Por el contrario, si A es no-singular (A equivalente por filas a I_n), $U = I_n$ y entonces

$$\det(U) = 1 \neq 0$$

que implica $\det(A) \neq 0$. ■

Segundo resultado clave:

2. Si A y B son matrices de $n \times n$:

$$\det(AB) = \det(A)\det(B)$$

Demostración. Si B es singular ($\det(B) = 0$), el sistema $B\mathbf{x}$ tiene soluciónes no triviales $\mathbf{x} \neq \mathbf{0}$, que son también solución de $(AB)\mathbf{x} = \mathbf{0}$, ya que $(AB)\mathbf{x} = A(B\mathbf{x}) = A\mathbf{0} = \mathbf{0}$. Esto muestra que AB es también singular. Por lo tanto, se cumple

$$\det(AB) = 0 = \det(A)\det(B)$$

Si B es no-singular, B se puede escribir como un producto de matrices elementales: $B = E_k \dots E_1$, y entonces

$$\begin{aligned}\det(AB) &= \det(AE_kE_{k-1}...E_1)\\ &= \det(A)\det(E_k).\det(E_{k-1})...\det(E_1)\\ &= \det(A)\det(E_kE_{k-1}...E_1)\\ &= \det(A)\det(B)\end{aligned}$$ ■

Importante: Una primer consecuencia de la propiedad anterior es la siguiente:
Si A es no singular, el determinante de la inversa A^{-1} es la inversa del determinante:

$$\det(A^{-1}) = (\det(A))^{-1}$$

Puede el lector probar este resultado a partir del determinante de un producto de matrices ($AA^{-1} = I$).

Ejemplo 3.5.1 Si

$$A = \begin{pmatrix}1 & 1\\ 2 & 1\end{pmatrix}, \qquad B = \begin{pmatrix}2 & 3\\ 1 & -1\end{pmatrix} \qquad \Rightarrow \qquad AB = \begin{pmatrix}3 & 2\\ 5 & 5\end{pmatrix}$$

y se verifica $\det(AB) = 5 = (-1)(-5) = \det(A)\det(B)$. Además,

$$\det(A^{-1}) = \begin{vmatrix}-1 & 1\\ 2 & -1\end{vmatrix} = -1 = \frac{1}{\det(A)}$$

$$\det(B^{-1}) = \det\left(\frac{1}{5}\begin{pmatrix}1 & 3\\ 1 & -2\end{pmatrix}\right) = \frac{-5}{25} = \frac{1}{-5} = \frac{1}{\det(B)}$$

Problemas 3.5

1. Indique mediante el determinante si las siguientes matrices son singulares o no singulares.

(a) $\begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix}$ **(b)** $\begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}$ **(c)** $\begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix}$

(d) $\begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 2 \\ 0 & 1 & 4 \end{pmatrix}$ **(e)** $\begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 4 & 1 & 3 \end{pmatrix}$ **(f)** $\begin{pmatrix} 2 & 1 & 1 \\ 2 & -2 & 1 \\ 1 & 0 & k \end{pmatrix}$

2. Cada par de matrices difiere por una operación elemental de fila. Usar esta operación para comparar $\det(A)$ con $\det(B)$.

(a) $A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}$

(b) $A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$

(c) $A = \begin{pmatrix} 1 & -1 & 3 \\ 2 & 2 & -6 \\ 1 & 0 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -1 & 3 \\ 1 & 1 & -3 \\ 1 & 0 & 4 \end{pmatrix}$

3. Dada la matriz

$$U = \begin{pmatrix} 4 & 4 & 8 & 8 \\ 0 & 1 & 2 & 2 \\ 0 & 1 & 2 & 6 \\ 4 & 1 & 1 & 2 \end{pmatrix}$$

calcular $\det(U)$, $\det(-3U^T)$ y $\det(U^{-1})$ (se sugiere llevarla a la forma triangular y luego usar propiedades).

4. a) Decidir si los siguientes sistemas tienen solución única, usando ahora el determinante de la matriz involucrada. En los casos de 3×3 calcular el determinante mediante la expansión por cofactores y mediante el procedimiento por operaciones elementales que lleva a una matriz triangular.

(a) $\begin{aligned} 3x + 6y &= 18 \\ x + 2y &= 6 \end{aligned}$ **(b)** $\begin{aligned} x + y &= 1 \\ x - y &= -1 \end{aligned}$

(c) $\begin{aligned} x_1 \phantom{{}-x_2} + x_3 &= 4 \\ x_1 - x_2 + 2x_3 &= 5 \\ 4x_1 - x_2 + 5x_3 &= 17 \end{aligned}$ **(d)** $\begin{aligned} x_1 + x_2 + 2x_3 &= 4 \\ x_1 - x_2 - x_3 &= 2 \\ 4x_1 + 2x_2 + rx_3 &= b \end{aligned}$

b) Resolver los sistemas anteriores, indicando el conjunto solución. En (d) determinar los valores de r y b para los que el sistema será compatible determinado, compatible indeterminado e incompatible, indicando el conjunto solución en los casos compatibles.

5. Sea A una matriz de $n \times n$.
a) Demostrar que $\det(A^k) = [\det(A)]^k$ para todo $k \geq 1 \in \mathbb{N}$.
b) Mostrar que si A es no singular, el resultado anterior vale también $\forall\ k$ entero

negativo.
c) Si se define $A^0 = I_n$, muestre que el resultado vale también para $k = 0$.

6. ¿Cuáles valores del número real x hacen que la matriz

$$A = \begin{pmatrix} 12-x & 4 \\ 8 & 8-x \end{pmatrix}$$

sea singular? Este problema aparecerá en el cálculo de autovalores de matrices.

7. a) ¿ Existen valores de θ para los que

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

es singular? Interpretar geométricamente.
b) Idem a) para $B = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.

8. Si A, B, C son todas matrices de $n \times n$,
a) ¿Es cierto que $\det(AB) = \det(BA)$ aún si $AB \neq BA$?
b) ¿Es cierto que $\det(ABC) = \det(BAC)$ aún si $ABC \neq BAC$?

9. Si $\det(A) = 2$ y $\det(B) = 3$, con A, B de $n\times n$, determinar $\det[-2A^2B^T(A^T)^{-1}B^{-2}]$.

10. a) Si A es una matriz singular de $n \times n$ y B es una matriz arbitraria de $n \times n$,
a) Muestre que AB, BA y BAB son singulares.
b) Muestre que $A + AB$, $A + BA$ y $A^2 + 2AB + A$ son singulares.

11. a) Muestre que si A es una matriz real ortogonal $(AA^T = I) \Rightarrow \det(A) = \pm 1$.
b) Si $\det(A) = \pm 1$, ¿puede afirmarse que A es ortogonal? (piense un posible contra-ejemplo).

12. a) Interprete geométricamente la identidad $\det(\alpha A) = \alpha^3 \det(A)$ para matrices A de 3×3.
b) Idem para $\det(B) = \alpha \det(A)$ si B se obtiene de A multiplicando una de las filas de A por α. (Considere como aumenta el volumen de un cubo si a) todas las aristas son aumentadas en un factor α o b) si una sola arista esa aumentada en tal factor).

3.6. Métodos para calcular el determinante

Hemos visto dos formas para calcular el determinante de una matriz A:

a) Usando la definición en términos de expansión de cofactores.

b) Reduciendo la matriz a una forma triangular (eliminación Gaussiana). En este caso, sólo es necesario "contar" la cantidad de intercambios de filas durante el proceso si se usan únicamente las operaciones elementales de tipo (I) y (III) (¿por qué?).

Cada término en la expansión de cofactores es un producto de n coeficientes de A, elegidos de tal manera que haya uno de cada columna y uno de cada fila. Esto es,

$$a_{1j_1}a_{2j_2}a_{3j_3}...a_{nj_n} \tag{3.20}$$

donde $\{j_1, j_2, \ldots, j_n\}$ es alguna permutación de los enteros $\{1, 2, \ldots, n\}$.
Por tanto, existen $n! = n(n-1)(n-2)\ldots 1$ posibilidades distintas de asignar las columnas a $\{j_1, j_2, \ldots, j_n\}$. Luego, existen $n!$ sumandos en total, todos de la forma (3.20), en la expresión completa del determinante de A, tal como se indicó en (3.9).
Entonces, para calcular $\det(A)$ por la fórmula que combina los cofactores, método (a), la cantidad de sumas necesarias es $n! - 1$. La cantidad de productos, usando el desarrollo por cofactores respecto de una fila o columna, al haber n productos de un elemento de una fila por el cofactor c_{ij} respectivo, es n[número de productos en un determinante de $(n-1)\times(n-1)]+n$. Esto es del orden de $n!$ ($O(n!)$) (aprox. $\approx (e-1)n!$ para n grande).

Por otro lado, puede verse que por el método de reducción a una matriz triangular (o métodos equivalentes) son necesarios esencialmente $n^3/3$ operaciones ($O(n^3)$) para el cálculo del determinante. En efecto, al dejar en 0 los elementos de la columna 1 por debajo del pivote a_{11}, mediante operaciones $f_i - \frac{a_{i1}}{a_{11}} f_1$ para $i = 2, \ldots, n$, se realizan $(n-1)\times(n-1)$ sumas y $(n-1)\times(n-1)+n-1 = n(n-1)$ productos. Por lo tanto, hasta llegar a la forma triangular el número todal de sumas es $\sum_{i=1}^{n}(i-1)^2 = (n(n-1)(2n-1))/6$ ($\approx n^3/3$ para n grande) y el número total de productos $\sum_{i=1}^{n}(i-1)i = n(n^2-1)/3$. El determinante requiere al final $n-1$ productos adicionales, por lo que el número total de productos es $n(n^2-1)/3 + n - 1$ (también $\approx n^3/3$ para n grande).

Como $n^3 \ll n!$ para n grande, el método de reducción por filas b) es númericamente mucho más eficiente que a). No obstante, el método a) permite obtener una expresión analítica del determinante. Esto es útil para determinar propiedades formales y también cuando los elementos de la matriz contienen parámetros arbitrarios o no conocidos (como sucede, como veremos más adelante, en el cálculo de autovalores).

Comparación entre el número de operaciones aritméticas necesarias para calcular el determinante de una matriz de $n \times n$, según los dos métodos anteriores.

	Expansión de cofactores (Método a)		Eliminación Gaussiana (Método b)	
n	Sumas	Multiplicaciones	Sumas	Multiplicaciones
2	1	2	1	3
3	5	9	5	10
4	23	40	14	23
5	119	205	30	44
10	3.628.799	6.235.300	285	339
20	$2{,}4 \times 10^{18}$	$4{,}2 \times 10^{18}$	2470	2679
$n \gg 1$	$n! - 1$	$\approx (e-1)n!$	$\approx n^3/3$	$\approx n^3/3$

Puede observarse que si $n > 3$, el método (b) es más eficiente (a menos que A tenga una cantidad significativa de ceros). Por esa razón, es el utilizado por los programas de cómputo corrientes.

Observación. Si A es singular, $\det(A) = 0$. Pero si $\det(A)$ es evaluado numéricamente (usando aritmética de computadora y no aritmética exacta), los errores de redondeo pueden ocasionar que el resultado no sea 0, aunque esté bastante cerca a 0.

Por lo tanto, en algunos casos es virtualmente imposible determinar computacionalmente cuando una matriz de $n \times n$ es verdaderamente singular. Se discutirá este aspecto con más detalle en la parte II.

Problema 3.6.1
Dado el determinante

$$\begin{vmatrix} 2 & 1 & 3 \\ 4 & 2 & 1 \\ 6 & -3 & 4 \end{vmatrix}$$

verifique que la cantidad de sumas y productos que se debe realizar para evaluarlo mediante el método (a) es 5 y 9, mientras que por el método (b) es 5 y 10 respectivamente. Verifique también que el valor del determinante es -60.

3.7. Matrices definidas por bloques

Las matrices definidas por bloques surgen frecuentemente en diversas aplicaciones. Daremos sus propiedades en forma de problemas.

1. Pruebe que el determinante de una matriz de la forma

$$M = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

donde A es una matriz de $n \times n$, B una matriz de $m \times m$ ($n \geq 1$, $m \geq 1$) y 0 denota matrices nulas (tal que M es de $(n+m) \times (n+m)$), es

$$\det(M) = \det(A)\det(B) \tag{3.21}$$

(Sugerencia: Demuestre primero (3.21) para una matriz A de 1×1, mediante el desarrollo por cofactores por la primer columna. Para el caso general, considere operaciones elementales que lleven A a una matriz triangular superior y aplique luego el resultado previo. Notar que la igualdad (3.21) es obvia si A y B son ambas triangulares superiores (¿Por qué ?). La igualdad puede también demostrarse escribiendo M como un producto conveniente, como se discute abajo).
Veremos en capítulos posteriores aplicaciones importantes de la propiedad (3.21).

2. Utilizando (3.21) evaluar el determinante de las matrices

$$M_1 = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 3 & 2 \\ 0 & 0 & 2 & 3 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 1 & 2 & 0 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 1 & 3 \end{pmatrix}$$

Verificar los resultados evaluando el determinante por otro método.

3. Muestre que es posible generalizar el resultado (3.21) a

$$\det\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \det\begin{pmatrix} A & 0 \\ D & B \end{pmatrix} = \det(A)\det(B) \qquad (3.22)$$

donde C es una matriz de $n \times m$ y D de $m \times n$, con A de $n \times n$, B de $m \times m$.

4. No obstante, mediante un contraejemplo muestre que en general,

$$\det\begin{pmatrix} A & C \\ D & B \end{pmatrix} \neq \det(A)\det(B) - \det(C)\det(D)$$

5. En cambio, si A de $n \times n$ es no singular, con B de $m \times m$, C de $n \times m$ y D de $m \times n$, muestre que

$$\det\begin{pmatrix} A & C \\ D & B \end{pmatrix} = \det(A)\det(B - DA^{-1}C) \qquad (3.23)$$

(Sugerencia: Muestre primero que $\begin{pmatrix} A & C \\ D & B \end{pmatrix} = \begin{pmatrix} A & 0 \\ D & I_m \end{pmatrix}\begin{pmatrix} I_n & A^{-1}C \\ 0 & B - DA^{-1}C \end{pmatrix}$ y use luego resultados anteriores).

6. Evaluar en base a los resultados anteriores los determinantes de las matrices

$$M_3 = \begin{pmatrix} 2 & 1 & 1 & 2 \\ 1 & 2 & 2 & 3 \\ 0 & 0 & 3 & 2 \\ 0 & 0 & 2 & 3 \end{pmatrix}, \quad M_4 = \begin{pmatrix} 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 2 \\ 3 & 2 & 0 & 0 \\ 2 & 3 & 0 & 0 \end{pmatrix}.$$

3.8. Regla de Cramer e inversa de una matriz

Presentaremos aquí expresiones **analíticas** para la inversa de una matriz A no singular y para la solución única del sistema lineal asociado $A\mathbf{x} = \boldsymbol{b}$, por medio de determinantes.

Inversa
Si A de $n \times n$ es no singular, entonces

$$A^{-1} = \frac{1}{\det(A)} C^T \tag{3.24}$$

donde C^T es la traspuesta de la matriz de cofactores C, de elementos

$$c_{ij} = (-1)^{i+j} \det(M_{ij})$$

y M_{ij} es la submatriz obtenida al suprimir la fila i y columna j de A. Es decir,

$$(A^{-1})_{ij} = \frac{c_{ji}}{\det(A)} = \frac{(-1)^{i+j} \det(M_{ji})}{\det(A)}$$

Demostración. Recordando la definición de determinante por cofactores, tenemos

$$(AC^T)_{ij} = \sum_{k=1}^{n} a_{ik} (C^T)_{kj} = \sum_{k=1}^{n} a_{ik} c_{jk} = \begin{cases} \det(A)\,, & i = j \\ 0 & i \neq j \end{cases}$$

ya que $\sum_{k=1}^{n} a_{ik} c_{jk}$ es, si $i = j$, el determinante de A desarrollado por la fila i, mientras que si $i \neq j$, es el determinante de una matriz similar a A pero con la fila j reemplazada por la fila i (dos filas iguales), que es nulo. Por lo tanto, si I_n es la matriz identidad,

$$AC^T = \det(A) I_n$$

Si $\det(A) \neq 0$, dividiendo por $\det(A)$ se obtiene entonces el resultado (3.24).

Regla de Cramer
Dado el sistema $A\mathbf{x} = \boldsymbol{b}$, con A de $n \times n$ **no singular**,

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

los elementos x_i del vector solución $\mathbf{x} = A^{-1}\boldsymbol{b}$ pueden expresarse como

$$x_i = \frac{\det A_i}{\det A} = \frac{1}{\det A} \begin{vmatrix} a_{11} & \dots & b_1 & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & b_n & \dots & a_{nn} \end{vmatrix} \tag{3.25}$$

donde A_i es la matriz obtenida al reemplazar la columna i de A por el vector columna $\boldsymbol{b}$.
Demostración. Aplicando la expresión (3.24) para A^{-1}, obtenemos, en forma explícita,

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A^{-1}\boldsymbol{b} = \frac{1}{\operatorname{der} A} C^T \boldsymbol{b} = \frac{1}{\det A} \begin{pmatrix} c_{11} & \dots & c_{n1} \\ \vdots & \ddots & \vdots \\ c_{1n} & \dots & c_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \frac{1}{\det A} \begin{pmatrix} c_{11}b_1 + \ldots + c_{n1}b_n \\ \vdots \\ c_{1n}b_1 + \ldots + c_{nn}b_n \end{pmatrix}$$

La fila i es entonces $x_i = \frac{1}{\det A}(b_1 c_{1i} + \ldots + b_n c_{ni}) = \frac{\det A_i}{\det A}$, ya que la suma $b_1 c_{1i} + \ldots + b_n c_{ni}$ es el desarrollo de $\det A_i$ por la columna i.

Observación: Estas expresiones proporcionan una expresión "analítica" para la inversa A^{-1} y la solución del sistema lineal asociado, en términos de determinantes. Resultan útiles para obtener propiedades generales de la solución y su dependencia con los elementos de la matriz A y del vector $\boldsymbol{b}$.

Por ejemplo, si $\det(A) = \pm 1$ y todos los elementos de A son enteros, muestra que **todos los elementos de la inversa son también enteros**.

No obstante, desde el punto de vista numérico no constituyen un método eficiente para resolver problemas con matrices numéricas de gran escala.

Ejemplo 3.8.1 Aplicando este método para A de 2×2, se obtiene directamente la expresión derivada en el capítulo de matrices:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Rightarrow C = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$$

Por lo tanto, si $\det(A) = ad - bc \neq 0$,

$$A^{-1} = \frac{1}{\det A} C^T = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Además, dado el sistema

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

aplicando (3.25) se obtiene

$$x = \frac{\begin{vmatrix} b_1 & b \\ b_2 & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}, \quad y = \frac{\begin{vmatrix} a & b_1 \\ c & b_2 \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}.$$

Ejemplo 3.8.2 Aplicando este método para A de 3×3, se obtiene

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \Rightarrow C = \begin{pmatrix} \begin{vmatrix} e & f \\ h & i \end{vmatrix} & -\begin{vmatrix} d & f \\ g & i \end{vmatrix} & \begin{vmatrix} d & e \\ g & h \end{vmatrix} \\ -\begin{vmatrix} b & c \\ h & i \end{vmatrix} & \begin{vmatrix} a & c \\ g & i \end{vmatrix} & -\begin{vmatrix} a & b \\ g & h \end{vmatrix} \\ \begin{vmatrix} b & c \\ e & f \end{vmatrix} & -\begin{vmatrix} a & c \\ d & f \end{vmatrix} & \begin{vmatrix} a & b \\ d & e \end{vmatrix} \end{pmatrix}$$

Por lo tanto, si $\det(A) \neq 0$,

$$A^{-1} = \frac{1}{\det A} C^T = \frac{1}{\det A} \begin{pmatrix} \begin{vmatrix} e & f \\ h & i \end{vmatrix} & -\begin{vmatrix} b & c \\ h & i \end{vmatrix} & \begin{vmatrix} b & c \\ e & f \end{vmatrix} \\ -\begin{vmatrix} d & f \\ g & i \end{vmatrix} & \begin{vmatrix} a & c \\ g & i \end{vmatrix} & -\begin{vmatrix} a & c \\ d & f \end{vmatrix} \\ \begin{vmatrix} d & e \\ g & h \end{vmatrix} & -\begin{vmatrix} a & b \\ g & h \end{vmatrix} & \begin{vmatrix} a & b \\ d & e \end{vmatrix} \end{pmatrix}$$

y la única solución del sistema

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

es

$$x = \frac{1}{\det A} \begin{vmatrix} b_1 & b & c \\ b_2 & e & f \\ b_3 & h & i \end{vmatrix}, \quad y = \frac{1}{\det A} \begin{vmatrix} a & b_1 & c \\ d & b_2 & f \\ g & b_3 & i \end{vmatrix}, \quad z = \frac{1}{\det A} \begin{vmatrix} a & b & b_1 \\ d & e & b_2 \\ g & h & b_3 \end{vmatrix}$$

Ejemplo 3.8.3

Consideremos el sistema

$$\begin{cases} x_1 + 0x_2 + 2x_3 & = & 6 \\ -3x_1 + 4x_2 + 6x_3 & = & 30 \\ -x_1 - 2x_2 + 3x_3 & = & 8 \end{cases}$$

Entonces $A = \begin{pmatrix} 1 & 0 & 2 \\ -3 & 4 & 6 \\ -1 & -2 & 3 \end{pmatrix}$, con $\det A = 44$ y obtenemos $A^{-1} = \frac{1}{44} \begin{pmatrix} 24 & -4 & -8 \\ 3 & 5 & -12 \\ 10 & 2 & 4 \end{pmatrix}$,

$$x_1 = \frac{\begin{vmatrix} 6 & 0 & 2 \\ 30 & 4 & 6 \\ 8 & -2 & 3 \end{vmatrix}}{\det(A)} = \frac{-10}{11}, \quad x_2 = \frac{\begin{vmatrix} 1 & 6 & 2 \\ -3 & 30 & 6 \\ -1 & 8 & 3 \end{vmatrix}}{\det(A)} = \frac{18}{11}, \quad x_3 = \frac{\begin{vmatrix} 1 & 0 & 6 \\ -3 & 4 & 30 \\ -1 & -2 & 8 \end{vmatrix}}{\det(A)} = \frac{38}{11}$$

Problemas 3.8

1. Muestre a partir de la regla de Cramer, que el elemento x_i de la solución del sistema $A\mathbf{x} = \boldsymbol{b}$ satisface

   $$\frac{\partial x_i}{\partial b_j} = \frac{c_{ji}}{\det(A)}$$

   con c_{ji} el cofactor j, i de A. Esta expresión determina la *variación* de los elementos de la solución con los parámetros independientes b_j.

2. Para un triángulo de lados a, b y c con ángulos opuestos α, β y γ respectivamente,
   a) Verificar usando trigonometría que

   $$b\cos(\gamma) + c\cos(\beta) = a$$

   $$c\cos(\alpha) + a\cos(\gamma) = b$$

   $$a\cos(\beta) + b\cos(\alpha) = c$$

   b) Aplicar la regla de Cramer para demostrar que

   $$\cos(\alpha) = \frac{b^2 + c^2 - a^2}{2bc}$$

   c) Obtener las expresiones de $\cos(\beta)$ y $\cos(\gamma)$

Capítulo 4

Espacios Vectoriales

4.1. Introducción

En este capítulo generalizaremos el concepto de vector y de espacio. Extenderemos las conocidas operaciones de suma de vectores y de multiplicación de un vector por un número real, ya vistas para vectores del plano ($\mathbb{R}^2$) o del espacio tridimensional ($\mathbb{R}^3$), a espacios de mayor dimensión y a conjuntos cuyos elementos no sean necesariamente pares o ternas de números reales. Por ejemplo, los elementos podrían ser matrices, polinomios, funciones, soluciones de ecuaciones lineales homogéneas, etc.

La idea central es definir un espacio vectorial como un conjunto de elementos (que se llamarán vectores) que tenga definidas dos operaciones básicas:

I. **La suma**

II. **La multiplicación por un escalar**

Estas dos operaciones deberán ser **cerradas** en el conjunto, es decir, dar como resultado otro vector del conjunto, y satisfacer ciertas propiedades que detallaremos a continuación.

Las operaciones anteriores permitirán definir la **combinación lineal** de vectores, que será también un vector del conjunto, y de esta forma generar los vectores mediante la combinación lineal de un cierto subconjunto de ellos. Esto posibilita una fácil caracterización de los elementos de espacios abstractos y a la vez interpretar los mismos geométricamente, mediante analogías con vectores de $\mathbb{R}^2$, $\mathbb{R}^3$ o en general $\mathbb{R}^n$. En particular, lograremos una comprensión más profunda de las soluciones de sistemas de ecuaciones lineales. Los espacios vectoriales abstractos juegan además un rol fundamental en la teoría de ecuaciones diferenciales lineales y en la física cuántica.

Ejemplo: $\mathbb{R}^2$

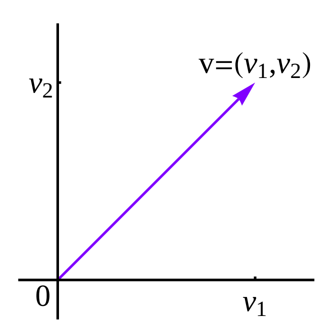


Figura 4.1: Vector en el plano

A modo de repaso, consideremos primero el conjunto $\mathbb{R}^2$ de pares ordenados de números reales $\boldsymbol{v} = (v_1, v_2)$ dotado de las operaciones:

- **Suma:**
  $\boldsymbol{u} + \boldsymbol{v} = (u_1, u_2) + (v_1, v_2) = (u_1 + v_1, u_2 + v_2)$

- **Multiplicación por un escalar (número real):**
  $\alpha\boldsymbol{v} = \alpha(v_1, v_2) = (\alpha v_1, \alpha v_2)$

Vemos que tanto la suma de dos vectores cualesquiera u, v de $\mathbf{R}^2$ como la multiplicación de cualquier vector v de $\mathbf{R}^2$ por cualquier número real α, da como resultado un vector de $\mathbf{R}^2$. Por lo cual decimos que el conjunto $\mathbf{R}^2$ de vectores del plano es cerrado bajo la operación de suma de vectores y bajo la m ultiplicación por un escalar real.

Geométricamente, $\mathbb{R}^2$ puede ser representado como el conjunto de todos los puntos del plano bidimensional. Un vector $\boldsymbol{v} = (v_1, v_2)$ puede ser representado como un segmento recto dirigido desde el vector nulo $\mathbf{0} = (0, 0)$ hasta (v_1, v_2).

El vector suma puede así obtenerse geométricamente mediante la conocida regla del paralelogramo, mientras que la multiplicación por un escalar α genera un vector con la misma dirección que el original, con el mismo sentido si $\alpha > 0$ (en la figura se ha elegido $\alpha > 1$) y el sentido opuesto si $\alpha < 0$.

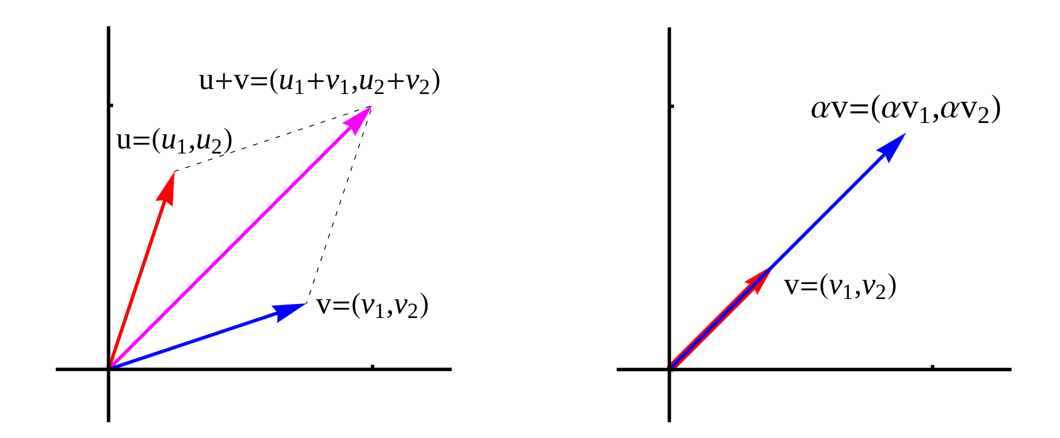


Figura 4.2: Suma de vectores y producto de un vector por un escalar

Estas operaciones verifican además las siguientes propiedades:

1. La suma es conmutativa: $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$

2. La suma es asociativa: $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$

3. Existe el vector nulo $\mathbf{0} = (0, 0)$ tal que $\boldsymbol{v} + \mathbf{0} = \boldsymbol{v} \ \forall \ \boldsymbol{v}$

4. Para todo $\boldsymbol{v} = (v_1, v_2)$ existe el vector opuesto $-\boldsymbol{v} = (-v_1, -v_2)$ tal que $\boldsymbol{v} + (-\boldsymbol{v}) = \mathbf{0}$

5. El producto por un escalar es distributivo respecto de la suma de vectores:
$\alpha(\boldsymbol{u} + \boldsymbol{v}) = \alpha\boldsymbol{u} + \alpha\boldsymbol{v}$

6. El producto por un escalar es distributivo respecto de la suma de escalares:
$(\alpha + \beta)\boldsymbol{v} = \alpha\boldsymbol{v} + \beta\boldsymbol{v}$

7. El producto por un escalar es asociativo: $\alpha\,(\beta\boldsymbol{v}) = (\alpha\beta)\boldsymbol{v}$

8. El producto por 1 no modifica el vector: $1\boldsymbol{v} = \boldsymbol{v} \ \forall \ \boldsymbol{v}$

Las mismas propiedades son satisfechas por el conjunto $\mathbb{R}^3$ de vectores en el espacio tridimensional, y en general $\mathbb{R}^n$.

A continuación extenderemos estas propiedades a conjuntos más generales. La idea es definir una estructura algebraica general, tal que cuando se pueda probar una propiedad para dichos conjuntos, la misma sea válida independientemente de los elementos que constituyan el conjunto.

4.2. Espacio vectorial

Definición
Un conjunto V dotado de dos operaciones cerradas:
I. La suma de elementos de V
II. La multiplicación de un elemento de V por un escalar
es un ***espacio vectorial*** siempre y cuando se cumplan las siguientes propiedades:

1. La suma es conmutativa: $\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u} \quad \forall\, \boldsymbol{u},\boldsymbol{v}\in V$

2. La suma es asociativa: $(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}=\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w}) \quad \forall\, \boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\in V$

3. Existe un único vector nulo $\mathbf{0}$ tal que $\boldsymbol{v}+\mathbf{0}=\boldsymbol{v}\ \forall\ \boldsymbol{v}\in V$

4. $\forall\ \boldsymbol{v}\in V$ existe el vector opuesto $-\boldsymbol{v}$ tal que $\boldsymbol{v}+(-\boldsymbol{v})=\mathbf{0}$

5. $\alpha(\boldsymbol{u}+\boldsymbol{v})=\alpha\boldsymbol{u}+\alpha\boldsymbol{v}\ \ \forall\ \boldsymbol{u},\boldsymbol{v}\in V$ y $\forall$ escalar α

6. $(\alpha+\beta)\boldsymbol{v}=\alpha\boldsymbol{v}+\beta\boldsymbol{v}\ \ \forall\ \boldsymbol{v}\in V$ y $\forall$ escalar α,β

7. $\alpha(\beta\boldsymbol{v})=(\alpha\beta)\boldsymbol{v}\ \ \forall\ \boldsymbol{v}\in V$ y $\forall$ escalar α,β

8. $1\boldsymbol{v}=\boldsymbol{v}\ \ \forall\ \boldsymbol{v}\in V$

Los elementos del espacio vectorial V se denominan ***vectores***. Si los escalares son números reales se dice que V es un ***espacio vectorial real.***

Los escalares pueden ser también números complejos, en cuyo caso se dice que V es ***un espacio vectorial complejo***.

Observación 1. La definición de espacio vectorial no exige que exista un producto entre vectores. Volveremos sobre este tema más adelante.

Observación 2. El conjunto de los escalares puede ser también cualquier conjunto de números que forme un cuerpo, tal como el conjunto de números racionales. Un cuerpo es un conjunto que tiene definida la suma y multiplicación entre sus elementos, las cuales deben ser operaciones cerradas, conmutativas y asociativas, con validez de la propiedad distributiva respecto a la suma y existencia de 0 (elemento neutro para la suma), 1 (elemento neutro para la multiplicación) y elemento opuesto para la suma e inverso para la multiplicación (con excepción del 0). El conjunto de los números reales y el conjunto de los números complejos son también cuerpos.

Observación 3. Si bien en este curso utilizaremos la suma "usual" cuando consideremos vectores de $\mathbb{R}^n$ o matrices, en principio cualquier operación binaria $+ : V\times V\to V$ que satisfaga todas las propiedades anteriores (y por su puesto, que sea de utilidad en un cierto problema o contexto) puede ser considerada como una "suma" válida de vectores.

Ejemplos 4.2: Algunos espacios vectoriales reales.

1) $V = \mathbb{R}^n$. Es el conjunto de todas las n-uplas $(x_1, \ldots, x_n)$ de números reales:

$$\mathbb{R}^n = \{(x_1, \ldots, x_n),\ x_i \in \mathbb{R},\ i = 1, \ldots, n\}$$

Para $\boldsymbol{u} = (u_1, \ldots, u_n)$, $\boldsymbol{v} = (v_1, \ldots, v_n)$ y α real, la suma y la multiplicación por un escalar se definen como

$$\boldsymbol{u} + \boldsymbol{v} = (u_1 + v_1, \ldots, u_n + v_n)$$

$$\alpha \boldsymbol{v} = (\alpha v_1, \ldots, \alpha v_n)$$

El vector nulo es $\mathbf{0} = (0, \ldots, 0)$ y el vector opuesto a $\boldsymbol{v}$ es $-\boldsymbol{v} = (-v_1, \ldots, -v_n)$. Se comprueba fácilmente que se cumplen las 8 propiedades anteriores.

Casos particulares son $\mathbb{R}^1 = \mathbb{R}$ (el conjunto de todos los números reales),
$\mathbb{R}^2 = \{(x_1, x_2),\ x_1, x_2 \in \mathbb{R}\}$ el conjunto de vectores del plano, y
$\mathbb{R}^3 = \{(x_1, x_2, x_3), x_1, x_2, x_3 \in \mathbb{R}\}$ el conjunto de vectores del espacio tridimensional.
Frecuentemente resulta conveniente, especialmente cuando se trabaja con matrices, escribir los vectores de $\mathbb{R}^n$ como vectores columna $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ en lugar de vectores fila.

2) $V = \mathbb{R}^{m \times n}$. Es el conjunto de todas las matrices de $m \times n$ con elementos reales:

$$\mathbb{R}^{m \times n} = \left\{ A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ & \ddots & \\ a_{m1} & \ldots & a_{mn} \end{pmatrix},\ a_{ij} \in \mathbb{R}, i = 1, \ldots, m,\ j = 1, \ldots, n \right\}$$

En este caso los vectores son matrices de $m \times n$. La suma de matrices de $m \times n$ y la multiplicación de una matriz de $m \times n$ por un escalar real son operaciones cerradas en $\mathbb{R}^{m \times n}$, tal como se vió en el capítulo de matrices. Recordemos que dadas dos matrices A, $B \in \mathbb{R}^{m \times n}$, de elementos a_{ij}, b_{ij}, estas operaciones se definen como

$$(A + B)_{ij} = a_{ij} + b_{ij}$$

$$(\alpha A)_{ij} = \alpha a_{ij}$$

para cada elemento i, j, con $i = 1, \ldots, m$, $j = 1 \ldots, n$.

El **vector nulo** es en este caso la matriz nula de $m \times n$ ($0_{ij} = 0\ \forall\ i, j$), mientras que el vector opuesto de la matriz A es la matriz con todos los elementos cambiados de signo: $(-A)_{ij} = -a_{ij}\ \forall\ i, j$.
Se verifica fácilmente que se cumplen también las ocho propiedades anteriores.

Casos particulares son:
$\mathbb{R}^{n \times n}$: El conjunto de matrices cuadradas de $n \times n$
$\mathbb{R}^{1 \times n}$: El conjunto de matrices fila de $1 \times n$ (idéntico a $\mathbb{R}^n$)
$\mathbb{R}^{n \times 1}$: El conjunto de matrices columna de $n \times 1$ (también identificado con $\mathbb{R}^n$)

3) $V = C[a,b]$. Es el conjunto de las funciones reales continuas definidas en el intervalo cerrado $[a,b]$:

$$V = \{f : [a,b] \to \mathbb{R},\ f\,\text{continua en}\,[a,b]\}$$

En este caso los vectores son las **funciones** f. Definiendo la suma de dos funciones y la multiplicación de una función por un escalar real como

$$(f+g)(t) = f(t) + g(t)$$

$$(\alpha f)(t) = \alpha f(t)$$

$\forall t \in [a,b]$, se verifica fácilmente que estas operaciones son cerradas en V: Si f y g son funciones reales continuas en $[a,b]$, tanto su suma $f+g$ como αf son también funciones reales continuas en ese intervalo.

El vector nulo es la función nula $0(t) = 0\ \forall\ t \in [a,b]$, mientras que el vector opuesto a f es $-f$, definido por $(-f)(t) = -f(t)\ \forall\ t \in [a,b]$. Las 8 propiedades anteriores se verifican fácilmente.

El conjunto $\mathbb{R}^{[a,b]}$ de todas las funciones reales (continuas o no) con dominio $[a,b]$, $\mathbb{R}^{[a,b]} = \{f : [a,b] \to \mathbb{R}\}$ es también un espacio vectorial con las operaciones anteriores, que incluye al espacio $C[a,b]$.

4) $V = P_n$. Es el conjunto de todos los polinomios reales de grado menor o igual a n:

$$P_n = \{p(t) = a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n,\ a_i \in \mathbb{R},\ i = 0, \ldots, n\}$$

P_n es un subconjunto del conjunto de funciones reales continuas con dominio todo $\mathbb{R}$. En este caso los vectores son polinomios. Resulta obvio que la suma de dos polinomios p y $q \in P_n$ es otro polinomio $\in P_n$, y lo mismo sucede con la multiplicación por un escalar real: Si $q(t) = b_0 + b_1 t + \ldots + b_n t^n$,

$$(p+q)(t) = (a_0 + b_0) + (a_1 + b_1)t + \ldots + (a_n + b_n)t^n$$

$$(\alpha p)(t) = \alpha a_0 + \alpha a_1 t + \ldots + \alpha a_n t^n$$

El vector nulo es el polinomio nulo $0(t) = 0 + 0t + \ldots + 0t^n$ y el opuesto a $p(t)$ es $-p(t) = -a_0 - a_1 t - \ldots - a_n t^n$. Es fácil ver que se verifican también las 8 propiedades anteriores.

5) $V = \{0\}$. Es el conjunto formado por el número real 0. Es un ejemplo trivial de espacio vectorial: Dado que $0 + 0 = 0$ y $\alpha 0 = 0\ \forall\ \alpha$, las operaciones de suma y multiplicación por un escalar son trivialmente cerradas en V. Se verifican también las restantes propiedades.

Nótese, no obstante, que el conjunto $\{1\}$ no es un espacio vectorial, ya que la suma no es una operación cerrada en el mismo: $1 + 1 = 2 \notin \{1\}$. Tampoco lo es la multiplicación por un escalar arbitrario.

Mencionemos ahora algunas propiedades básicas válidas en todo espacio vectorial ($\mathbf{0}$ denota el vector nulo y 0 el escalar nulo):

Teorema 4.2.1
Sea V un espacio vectorial. Entonces:
a) Para todo escalar α, $\alpha\,\mathbf{0} = \mathbf{0}$
b) Para todo $\boldsymbol{v} \in V$, $0\,\boldsymbol{v} = \mathbf{0}$
c) Si $\alpha\,\boldsymbol{v} = \mathbf{0} \;\Rightarrow \alpha = 0$ o $\boldsymbol{v} = \mathbf{0}$ (o ambos nulos)
d) Para todo $\boldsymbol{v} \in V$, $(-1)\,\boldsymbol{v} = -\boldsymbol{v}$

Demostración.
a) $\alpha\,\mathbf{0} = \alpha\,(\mathbf{0}+\mathbf{0}) = \alpha\,\mathbf{0} + \alpha\,\mathbf{0}$, utilizando las propiedades **3.** y **5.** Sumando a ambos miembros de esta ecuación el opuesto $-\alpha\,\mathbf{0}$ y utilizando las propiedades **4.** y **2.** se obtiene: $\mathbf{0} = \alpha\,\mathbf{0} + \mathbf{0}$ y por lo tanto, usando **3.**, $\mathbf{0} = \alpha\,\mathbf{0}$.
b) La demostración es análoga a la de a), partiendo de $0 = 0 + 0$. Se deja como ejercicio.
c) Si $\alpha = 0$ ya fué probado en b).
Supongamos ahora $\alpha \neq 0$. Multiplicando a ambos miembros de $\alpha\mathbf{v} = \mathbf{0}$ por $^1/_\alpha$ y utilizando a) se obtiene: $(^1/_\alpha)(\alpha\boldsymbol{v}) = (^1/_\alpha)\mathbf{0} = \mathbf{0}$. Utilizando ahora **7.** y **8.**, $(^1/_\alpha)(\alpha\boldsymbol{v}) = ((^1/_\alpha)\alpha)\mathbf{v} = 1\boldsymbol{v} = \boldsymbol{v}$. Por lo tanto $\boldsymbol{v} = \mathbf{0}$.
d) Se deja también como ejercicio.

Combinaciones lineales de vectores

Sea V un espacio vectorial. Si $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ son vectores de V y α_1, α_2 escalares, entonces la suma

$$\alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2$$

se denomina **combinación lineal** de $\boldsymbol{v}_1$ y $\boldsymbol{v}_2$, y es siempre un vector de V.
Análogamente, si $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ son n vectores de V y $\alpha_1, \ldots, \alpha_n$ escalares, la suma

$$\alpha_1\boldsymbol{v}_1 + \ldots + \alpha_n\boldsymbol{v}_n$$

se denomina **combinación lineal** de los vectores $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ y es siempre un vector de V.

La demostración de que la combinación lineal de vectores es un vector del espacio es inmediata: Como V es cerrado bajo multiplicación por un escalar, tanto $\alpha_1\boldsymbol{v}_1$ como $\alpha_2\boldsymbol{v}_2$ son siempre vectores de V, para cualquier par de escalares α_1 y α_2. Y como V es también cerrado bajo la suma de vectores, entonces $\alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2$ es también un vector de V. La demostración del caso general con n vectores es similar.

Esto implica que un espacio vectorial contiene a toda combinación lineal de sus vectores. Además, veremos luego que en muchos casos es posible generar cualquier vector del espacio mediante la combinación lineal de un conjunto finito de vectores.

4.3. Subespacios

Un ***subespacio*** S de un espacio vectorial V es un subconjunto no vacío de V que es también un **espacio vectorial**. Esto implica que S debe ser cerrado bajo las operaciones de suma de vectores y de multiplicación por un escalar.

Como todos los elementos de S pertenecen a V, las ocho propiedades se satisfacen automáticamente, por lo cual para determinar si S es un subespacio bastará comprobar las condiciones de clausura de la suma y el producto. El vector nulo $\mathbf{0}$ de V deberá necesariamente pertenecer a S para que pueda cumplirse la clausura.

Resumiendo:

Un subconjunto S de un espacio vectorial V es un ***subespacio*** de V si se cumplen: **1.** $\mathbf{0} \in S$ (esto garantiza que S es no vacío) **2.** Si $\boldsymbol{u} \in S$ y $\boldsymbol{v} \in S \Rightarrow \boldsymbol{u} + \boldsymbol{v} \in S$ (S es cerrado con respecto a la suma) **3.** Si $\boldsymbol{u} \in S \Rightarrow \alpha.\boldsymbol{u} \in S$ $\forall$ escalar α (S es cerrado con respecto al producto por un escalar)

Nótese que el vector nulo $\{\mathbf{0}\}$ es siempre un subespacio de V (subespacio nulo).

Cualquier subespacio de V distinto de V y del subespacio nulo $\{\mathbf{0}\}$ se denomina ***subespacio propio*** de V.

Ejemplos 4.3

1) Sea $V = \mathbb{R}^2$ y S el subconjunto de $\mathbb{R}^2$ formado por vectores de la forma $(x, 0)$ con x real arbitrario, es decir,

$$S = \{(x, 0), x \in \mathbb{R}\}$$

Geométricamente S es el eje x.

S es un subespacio de $\mathbb{R}^2$ pues:

1. $\mathbf{0} = (0, 0) \in S$ (se obtiene para $x = 0$).

2. Si $\boldsymbol{v}_1 = (x_1, 0)$ y $\boldsymbol{v}_2 = (x_2, 0)$ son dos vectores de S,

$$\begin{aligned} \boldsymbol{v}_1 + \boldsymbol{v}_2 &= (x_1, 0) + (x_2, 0) \\ &= (x_1 + x_2, 0) \in S \;\; \text{(corresponde a } x = x_1 + x_2\text{)}. \end{aligned}$$

3. Si $\boldsymbol{v} = (x, 0) \in S$,

$$\begin{aligned} \alpha\boldsymbol{v} &= \alpha(x, 0) \\ &= (\alpha x, 0) \in S \end{aligned}$$

Al cumplirse **1.**, **2.** y **3.** podemos afirmar que S es un subespacio de $\mathbb{R}^2$.
Geométricamente, este resultado es obvio: la suma de dos vectores situados sobre el eje x es otro vector sobre el eje x, y al multiplicar cualquier vector sobre este eje por un escalar se obtiene un vector sobre este mismo eje. El eje y ($\{(0, y),\ y \in \mathbb{R}\}$) es obviamente también un subespacio de $\mathbb{R}^2$.

2) Sea $V = \mathbb{R}^2$ y C el subconjunto formado por vectores de la forma $(x, 1)$, es decir,

$$C = \{(x, 1), x \in \mathbb{R}\}$$

Geométricamente C es una recta horizontal que pasa por $(0, 1)$.

C **no** es un subespacio de $\mathbb{R}^2$ pues $\mathbf{0} = (0, 0) \notin S$. Esto ya basta para mostrarlo. Notemos también que C no es cerrado bajo la suma de vectores, ya que si $\boldsymbol{v}_1 = (x_1, 1)$ y $\boldsymbol{v}_2 = (x_2, 1)$ son vectores de $S \Rightarrow \boldsymbol{v}_1 + \boldsymbol{v}_2 = (x_1, 1) + (x_2, 1) = (x_1 + x_2, 2) \notin S$. C tampoco es cerrado bajo la multiplicación por un escalar, ya que si $\boldsymbol{v} = (x, 1) \in S$ $\Rightarrow \alpha\boldsymbol{v} = \alpha(x, 1) = (\alpha x, \alpha) \notin S$ para $\alpha \neq 1$.
C es en realidad un subespacio trasladado (denominado subespacio afin).

3) Sea $V = \mathbb{R}^2$ y S el subconjunto de $\mathbb{R}^2$ formado por vectores de la forma (x, mx):

$$S = \{(x, y) \in \mathbb{R}^2,\ y = mx\}$$

con m fijo. Geométricamente S es una recta con pendiente m que pasa por el origen.

S es un subespacio de $\mathbb{R}^2$ pues:
1. $\mathbf{0} = (0, 0) \in S$ (se obtiene para $x = 0$).
2. Si $\boldsymbol{v}_1 = (x_1, mx_1)$ y $\boldsymbol{v}_2 = (x_2, mx_2) \in S$,

$$\begin{aligned}
\boldsymbol{v}_1 + \boldsymbol{v}_2 &= (x_1, mx_1) + (x_2, mx_2) \\
&= (x_1 + x_2, mx_1 + mx_2) \\
&= (x_1 + x_2, m(x_1 + x_2)) \in S \text{ (corresponde a } x = x_1 + x_2)
\end{aligned}$$

3. Si $\boldsymbol{v} = (x, mx) \in S$

$$\begin{aligned}
\alpha\boldsymbol{v} &= \alpha(x, mx) \\
&= (\alpha x, \alpha mx) \\
&= (\alpha x, m(\alpha x)) \in S
\end{aligned}$$

Al cumplirse **1.**, **2.** y **3.**, podemos afirmar que S es un subespacio de $\mathbb{R}^2$. Geométricamente, es obvio que la suma de dos vectores pertenecientes a esta recta es otro vector sobre la misma recta, y que la multiplicación de estos vectores por un escalar también da como resultado un vector sobre la misma recta.

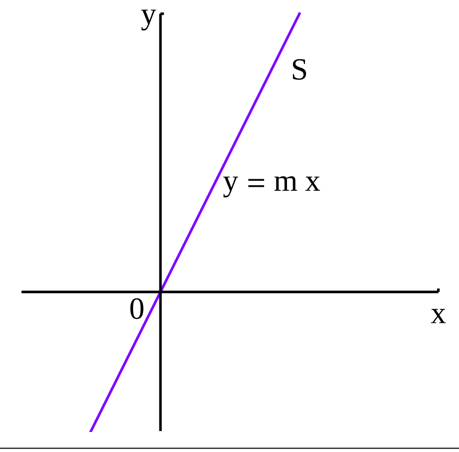


Figura 4.3: Todos los subespacios propios de $\mathbb{R}^2$ son **rectas que pasan por el origen**

4) Sea $V = \mathbb{R}^2$ y C el conjunto $C = \{(x, y) \in \mathbb{R}^2,\ y = mx + b,\ b \neq 0\}$, que geométricamente corresponde a una recta que no pasa por el origen. Dado que C no contiene al origen $\mathbf{0} = (0, 0)$, C **no** es un subespacio de $\mathbb{R}^2$ (es un subespacio trasladado o afin). Tampoco es cerrado bajo suma o multiplicación por escalar (¡probar!).

5) Sea $V = \mathbb{R}^2$ y C el semiplano superior,

$$C = \{(x, y), x, y \in \mathbb{R},\ y \geq 0\}$$

C **no** es un subespacio de $\mathbb{R}^2$: Si bien $\mathbf{0} = (0,\ 0) \in C$ y C es cerrado bajo suma de vectores (¡probar!), C no es cerrado bajo la multiplicación por un escalar:
Si $\boldsymbol{v} = (x,\ y)$ con $y > 0 \Rightarrow \alpha\boldsymbol{v} = (\alpha x,\ \alpha y) \notin C$ si $\alpha < 0$, ya que $\alpha y < 0$.
Por ejemplo, $(0,\ 1) \in C$ pero $-(0,\ 1) = (0, -1) \notin C$.

6) Sea $V = \mathbb{R}^3$ y $S = \{(x, y, 0), x, y \in \mathbb{R}\}$ el plano xy. Se deja como ejercicio probar que S es un subespacio de $\mathbb{R}^3$. Por otro lado, $C = \{(x,\ y,\ 1),\ x,\ y \in \mathbb{R}\}$ no es un subespacio de $\mathbb{R}^3$ (¡probar!).

7) Generalizando el caso anterior, sea $V = \mathbb{R}^3$ y

$$S = \{(x, y, z) \in \mathbb{R}^3, ax + by + cz = 0\}$$

Geométricamente S es un plano que pasa por el origen perpendicular al vector (a, b, c) (no nulo). S es subespacio de $\mathbb{R}^3$ pues:
1. $\mathbf{0} = (0, 0, 0) \in S$ (se obtiene para $x = y = z = 0$).
2. Si $\boldsymbol{v}_1 = (x_1, y_1, z_1)$ y $\boldsymbol{v}_2 = (x_2, y_2, z_2)$ son vectores de S ($ax_i + by_i + cz_i = 0$ para $i = 1, 2$) $\Rightarrow \boldsymbol{v}_1 + \boldsymbol{v}_2 = (x_1 + x_2, y_1 + y_2, z_1 + z_2) \in S$ pues

$$a(x_1 + x_2) + b(y_1 + y_2) + c(z_1 + z_2) = (ax_1 + by_1 + cz_1) + (ax_2 + by_2 + cz_2) = 0 + 0 = 0$$

3. Si $\boldsymbol{v} = (x, y, z) \in S$ ($ax + by + cz = 0$) $\Rightarrow \alpha\boldsymbol{v} = (\alpha x, \alpha y, \alpha z) \in S$ pues

$$a(\alpha x) + b(\alpha y) + c(\alpha z) = \alpha(ax + by + cz) = \alpha 0 = 0$$

Al cumplirse **1.**, **2.** y **3.** podemos afirmar que S es un subespacio de $\mathbb{R}^3$. Geométricamente el resultado es obvio: La suma de vectores de este plano y la multiplicación de ellos por un escalar no salen del plano.

8) Sea $V = \mathbb{R}^3$ y

$$S = \{t(a, b, c), t \in \mathbb{R}\}$$

Geométricamente S es una recta que pasa por el origen con vector director (a, b, c) (no nulo), perpendicular al plano anterior.

S es subespacio de $\mathbb{R}^3$ pues:
1. $\mathbf{0} = (0, 0, 0) \in S$ (se obtiene para $t = 0$).
2. Si $\boldsymbol{v}_1 = t_1(a, b, c)$ y $\boldsymbol{v}_2 = t_2(a, b, c)$ son vectores de $S \Rightarrow \boldsymbol{v}_1 + \boldsymbol{v}_2 = (t_1 + t_2)(a, b, c) \in S$ (corresponde a $t = t_1 + t_2$).
3. Si $\boldsymbol{v} = t(a, b, c) \in S \Rightarrow \alpha\boldsymbol{v} = (\alpha t)(a, b, c) \in S$ (corresponde a $t \to \alpha t$).

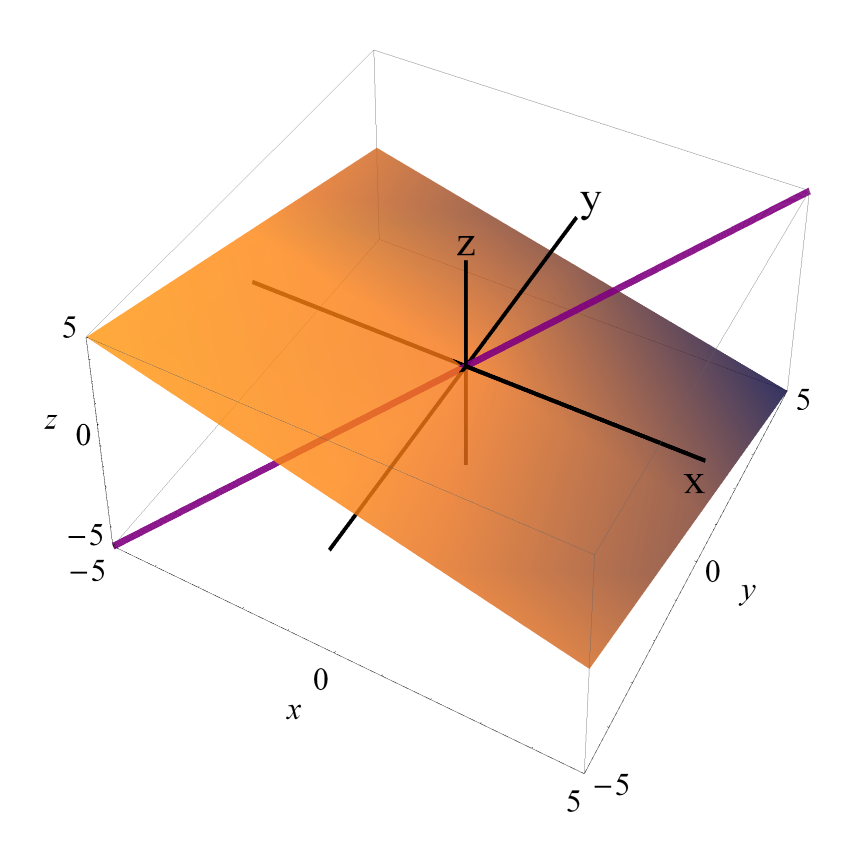


Figura 4.4: Todos los subespacios propios de $\mathbb{R}^3$ son **rectas o planos que pasan por el origen**

9) Sea $V = \mathbb{R}^{2\times 2}$ y S el conjunto de matrices de 2×2 de traza nula:

$$S = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}^{2\times 2}, a + d = 0 \right\}$$

S puede ser también escrito como $S = \left\{ \begin{pmatrix} a & b \\ c & -a \end{pmatrix}, a, b, c \in \mathbb{R} \right\}$.

S es subespacio de $\mathbb{R}^{2\times 2}$ pues:

1. La matriz nula $0 = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right) \in S$ (caso $a = b = c = 0$).

2. Si $A_1 = \left(\begin{smallmatrix} a_1 & b_1 \\ c_1 & -a_1 \end{smallmatrix}\right)$, $A_2 = \left(\begin{smallmatrix} a_2 & b_2 \\ c_2 & -a_2 \end{smallmatrix}\right)$ son dos matrices $\in S$, $A_1 + A_2 = \begin{pmatrix} a_1 + a_2 & b_1 + b_2 \\ c_1 + c_2 & -(a_1 + a_2) \end{pmatrix}$ $\in S$ ya que es también de traza nula.

3. Si $A \in S \Rightarrow \alpha A = \begin{pmatrix} \alpha a & \alpha b \\ \alpha c & -\alpha a \end{pmatrix} \in S$ ya que también es de traza nula.

Al cumplirse **1.**, **2.**, **3.** podemos afirmar que S es un subespacio de $\mathbb{R}^{2\times 2}$. Este resultado permanece válido para matrices de $n \times n$ y puede también demostrarse a partir de la linealidad de la operación de traza, como veremos en un capítulo posterior.

10) Sea $V = P_2 = \{p(t) = a_0 + a_1 t + a_2 t^2,\ a_i \in \mathbb{R}\}$, el espacio vectorial de los polinomios de grado menor o igual a 2 con coeficientes reales.

Veamos que $S = \{p(t) \in P_2, a_1 + a_2 = 0\}$ es un subespacio de P_2. Nótese que S es el subconjunto de los polinomios de P_2 de la forma $p(t) = a_0 + a_1(t - t^2)$, es decir, de los que satisfacen $p(0) = p(1)$.

1. $\mathbf{0} = 0 + 0t + 0t^2 \in S$, pues en este caso $a_1 + a_2 = 0 + 0 = 0$.
2. Sean $p_1(t) = a_0 + a_1 t + a_2 t^2 \in S$ y $p_2(t) = b_0 + b_1 t + b_2 t^2 \in S$. Es decir $a_1 + a_2 = 0$ y $b_1 + b_2 = 0$. Entonces
$p_1(t)+p_2(t) = (a_0 + a_1 t + a_2 t^2)+(b_0 + b_1 t + b_2 t^2) = (a_0+b_0)+(a_1+b_1)t+(a_2+b_2)t^2$ $\in S$ pues $(a_1 + b_1) + (a_2 + b_2) = (a_1 + a_2) + (b_1 + b_2) = 0 + 0 = 0$.
3. Sea $\alpha \in \mathbb{R}$ y $p(t) = a_0 + a_1 t + a_2 t^2 \in S$, es decir $a_0 + a_2 = 0$. Entonces
$\alpha p(t) = \alpha(a_0 + a_1 t + a_2 t^2) = (\alpha a_0) + (\alpha a_1)t + (\alpha a_2)t^2 \in S$, pues $(\alpha a_1) + (\alpha a_2) = \alpha(a_1 + a_2) = \alpha 0 = 0$.
Hemos entonces probado que S es un subespacio de P_2. Esto puede también demostrarse a partir de las otras formas de definir este subespacio, mencionadas arriba.

11) El producto escalar entre dos vectores $\boldsymbol{u} = (u_1, \ldots, u_n)$, $\boldsymbol{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n$ se define como $\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + \ldots + u_n v_n$. Y dos vectores son ortogonales si $\boldsymbol{u} \cdot \boldsymbol{v} = 0$. Mostraremos que el conjunto de vectores de $\mathbb{R}^n$ **ortogonales** a un vector dado $\boldsymbol{u}$, $S = \{\boldsymbol{v} \in \mathbb{R}^n, \boldsymbol{v} \cdot \boldsymbol{u} = 0\}$, es un subespacio de $\mathbb{R}^n$ (subespacio ortogonal a $\boldsymbol{u}$):

1. $\mathbf{0} \in S$ pues $\mathbf{0} \cdot \boldsymbol{u} = 0$
2. Si $\boldsymbol{v}_1$ y $\boldsymbol{v}_2 \in S$ $(\boldsymbol{v}_1 \cdot \boldsymbol{u} = 0,\ \boldsymbol{v}_2 \cdot \boldsymbol{u} = 0) \Rightarrow (\boldsymbol{v}_1 + \boldsymbol{v}_2) \cdot \boldsymbol{u} = \boldsymbol{v}_1 \cdot \boldsymbol{u} + \boldsymbol{v}_2 \cdot \boldsymbol{u} = 0 + 0 = 0$, por lo que $\boldsymbol{v}_1 + \boldsymbol{v}_2 \in S$
3. Si $\boldsymbol{v} \in S$ $(\boldsymbol{v} \cdot \boldsymbol{u} = 0) \Rightarrow (\alpha \boldsymbol{v}) \cdot \boldsymbol{u} = \alpha(\boldsymbol{v} \cdot \boldsymbol{u}) = \alpha 0 = 0$, por lo que $\alpha \boldsymbol{v} \in S$.
Por lo tanto S es un subespacio de $\mathbb{R}^n$. Si $\boldsymbol{u} = \mathbf{0} \Rightarrow S = \mathbb{R}^n$, pero si $\boldsymbol{u} \neq \mathbf{0}$, S será un subespacio propio de $\mathbb{R}^n$ (de dimensión $n - 1$, como veremos luego).

Se deja como ejercicio probar que el conjunto de vectores ortogonales a un cierto conjunto de vectores $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m\} \subset \mathbb{R}^n$ es también un subespacio de $\mathbb{R}^n$.

Problemas 4.3

1) Analizar si S es un subespacio del espacio vectorial indicado, e interpretar S geométricamente.
1.1) $V = \mathbb{R}^2$
a) $S = \{(x, y) \in \mathbb{R}^2,\ y = x\}$ b) $S = \{(x, y) \in \mathbb{R}^2,\ y = x^2\}$
c) $S = \{(x, y) \in \mathbb{R}^2,\ y \geq x\}$ d) $S = \{(x, y) \in \mathbb{R}^2,\ x^2 + y^2 \leq 1\}$
1.2) $V = \mathbb{R}^3$
a) $S = \{(x, y, z) \in \mathbb{R}^3,\ x + y + z = 0\}$ b) $S = \{(x, y, z) \in \mathbb{R}^3,\ x + y + z = -1\}$
c) $S = \{(x, y, z) \in \mathbb{R}^3,\ y = 0\}$ d) $S = \{(x, y, z) \in \mathbb{R}^3,\ y = 0, x = z\}$
e) $S = \{(x, y, z) \in \mathbb{R}^3,\ z \geq x^2 + y^2\}$ f) $S = \{(x, y, z) \in \mathbb{R}^3,\ x \neq 1\}$
1.3) $V = \mathbb{R}^4$ a) $S = \{(x, y, z, t) \in \mathbb{R}^4,\ x + y + z = t\}$

2) Probar que toda recta que pasa por el origen en $\mathbb{R}^2$ es un subespacio de $\mathbb{R}^2$. Mostrar también que las rectas que no pasan por el origen no son subespacios de $\mathbb{R}^2$.

3) Muestre que el conjunto de vectores de $\mathbb{R}^4$ ortogonales al vector $(1, 1, 1, 1)$ es un subespacio de $\mathbb{R}^4$.

4) Analice si el subconjunto S de matrices dado es un subespacio de $\mathbb{R}^{2\times 2}$.
a) $S = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}^{2\times 2}, b = c \right\}$ (conjunto de matrices simétricas de 2×2)
b) $S = \left\{ \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} \in \mathbb{R}^{2\times 2} \right\}$ (conjunto de matrices diagonales de 2×2)
c) $S = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}^{2\times 2}, ad - bc = 0 \right\}$ (conjunto de matrices singulares de 2×2)

5) Analice si el subconjunto S de matrices dado es un subespacio de $\mathbb{R}^{n\times n}$.
a) $S = \{A \in \mathbb{R}^{n\times n}, A^T = A\}$ (conjunto de matrices simétricas de $n \times n$)
b) $S = \{A \in \mathbb{R}^{n\times n}, A^T = -A\}$ (conjunto de matrices antisimétricas de $n \times n$)
c) $S = \{A \in \mathbb{R}^{n\times n}, a_{ij} = 0 \text{ si } i \neq j\}$ (conjunto de matrices diagonales de $n \times n$)
d) $S = \{A \in \mathbb{R}^{n\times n}, \det A = 0\}$ (conjunto de matrices singulares de $n \times n$)
e) $S = \{A \in \mathbb{R}^{n\times n}, \det A \neq 0\}$ (conjunto de matrices no-singulares de $n \times n$)
f) $S = \{A \in \mathbb{R}^{n\times n}, a_{ij} = 0 \text{ si } i > j\}$ (matrices triangulares superiores de $n \times n$)

6) Analice si el subconjunto de funciones $f : \mathbb{R} \to \mathbb{R}$ derivables $\forall\, x \in \mathbb{R}$ es un subespacio del espacio $C(\mathbb{R})$ de funciones $f : \mathbb{R} \to \mathbb{R}$ continuas.

7) a) Determine si $S = \{f : \mathbb{R} \to \mathbb{R}, \frac{df}{dx} - f = 0\}$ (el conjunto de funciones que satisfacen $\frac{df}{dx} = f$) es un subespacio del espacio de funciones continuas $C(\mathbb{R})$.
b) Idem para $S = \{f : \mathbb{R} \to \mathbb{R}, \frac{df}{dx} - f = 1\}$.

8) Sea $V = P_2$ el espacio de polinomios de grado ≤ 2. Determine si el subconjunto de polinomios de P_2 que satisface $p(1) = 0$ es un subespacio de P_2. ¿Sucede lo mismo con el conjunto de polinomios de P_2 que satisface $p(1) = 1$?

9) Sean S_1 y S_2 subespacios de un espacio vectorial V. Probar que la intersección $S_1 \bigcap S_2$ es un subespacio de V, y que la unión $S_1 \bigcup S_2$ no es necesariamente un subespacio de V.

4.4. Espacio nulo de una matriz

Sea A una matriz de $m \times n$. Definimos el ***espacio nulo*** de A, $N(A)$, como el conjunto de todas las soluciones del sistema lineal homogéneo $A\,\boldsymbol{v} = \mathbf{0}$:

$$N(A) = \{\boldsymbol{v} \in \mathbb{R}^n : A\,\boldsymbol{v} = \mathbf{0}\}$$

(donde $\mathbb{R}^n \equiv \mathbb{R}^{n\times 1}$ denota aquí el espacio de vectores columnas reales de $n \times 1$).
$N(A)$ es un **subespacio** de $\mathbb{R}^n$.

Veamos que es subespacio de $\mathbb{R}^n$:

1. $\mathbf{0} \in N(A)$, pues $A\,\mathbf{0} = \mathbf{0}$. El sistema homogéneo tiene siempre al menos la solución trivial $\boldsymbol{v} = \mathbf{0}$.

2. $N(A)$ es cerrado con respecto a la suma: Si $\boldsymbol{u}$, $\boldsymbol{v} \in N(A)$, por lo cual $A\,\boldsymbol{u} = \mathbf{0}$ y $A\,\boldsymbol{v} = \mathbf{0}$,
$$A\,(\boldsymbol{u} + \boldsymbol{v}) = A\,\boldsymbol{u} + A\,\boldsymbol{v} = \mathbf{0} + \mathbf{0} = \mathbf{0}$$
por lo que $\boldsymbol{u} + \boldsymbol{v} \in N(A)$.

3. $N(A)$ es cerrado con respecto a la multiplicación por escalar: Si $\alpha \in \mathbb{R}$ y $\boldsymbol{v} \in N(A)$,
$$A\,(\alpha\,\boldsymbol{v}) = \alpha\,(A\,\boldsymbol{v}) = \alpha\,\mathbf{0} = \mathbf{0}$$
por lo que $\alpha\boldsymbol{v} \in N(A)\ \forall\ \alpha$. Por lo tanto $N(A)$ es un subespacio de $\mathbb{R}^n$.

Interpretación geométrica. Dado que la fila i de $A\boldsymbol{v}$ es el producto escalar de la fila i de A por $\boldsymbol{v}$, es decir $\sum_j a_{ij}v_j$, el espacio nulo tiene una clara interpretación geométrica: Es el conjunto de vectores que son **ortogonales** (o sea perpendiculares) a **todas** las filas de la matriz A, es decir, es el **subespacio ortogonal** a todas las filas de A. Volveremos sobre este punto en la sección 4.11 (ver gráfico 4.9) y en la parte II.

Ejemplo 4.4.1 Sea $A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 2 & 3 & 1 & 2 \end{pmatrix}$

Entonces $N(A) = \left\{\boldsymbol{v} \in \mathbb{R}^4 : A.\boldsymbol{v} = \mathbf{0}\right\} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbb{R}^4 : \left\{ \begin{array}{rcl} x_1 + x_2 + x_4 & = & 0 \\ 2x_1 + 3x_2 + x_3 + 2x_4 & = & 0 \end{array} \right. \right\}$

Aplicando la reducción de Gauss-Jordan obtenemos

$$(A \mid \mathbf{0}) = \left(\begin{array}{cccc|c} 1 & 1 & 0 & 1 & 0 \\ 2 & 3 & 1 & 2 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{cccc|c} 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{cccc|c} 1 & 0 & -1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{array}\right)$$

Tomando como variables libres a x_3, x_4, se obtiene $x_1 = x_3 - x_4$, $x_2 = -x_3$. Entonces

$$N(A) = \left\{ \begin{pmatrix} x_3 - x_4 \\ -x_3 \\ x_3 \\ x_4 \end{pmatrix}, x_3, x_4 \in \mathbb{R} \right\} = \left\{ x_3 \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix},\ x_3, x_4 \in \mathbb{R} \right\}$$

que es un subespacio de $\mathbb{R}^4$ (como comprobaremos en breve). Puede verificar el lector que los dos vectores columna que generan $N(A)$ son ortogonales a todas las filas de A.

Sistemas no homogéneos. El conjunto de soluciones de un sistema no homogéneo $A\boldsymbol{v} = \boldsymbol{b}$ ($\boldsymbol{b} \neq \boldsymbol{0}$), **no** es un subespacio de $\mathbb{R}^n$, pues no contiene el vector nulo $\boldsymbol{0}$ y no es cerrado con respecto a la suma y al producto por un escalar: Si $\boldsymbol{u}$ y $\boldsymbol{v}$ son soluciones del sistema ($A\boldsymbol{u} = \boldsymbol{b}$, $A\boldsymbol{v} = \boldsymbol{b}$) $\Rightarrow A(\boldsymbol{u}+\boldsymbol{v}) = A\boldsymbol{u} + A\boldsymbol{v} = \boldsymbol{b}+\boldsymbol{b} = 2\boldsymbol{b}$ ($\neq \boldsymbol{b}$ si $\boldsymbol{b} \neq \boldsymbol{0}$) y $A(\alpha\,\boldsymbol{v}) = \alpha\,(A\boldsymbol{v}) = \alpha\,\boldsymbol{b}$ ($\neq \boldsymbol{b}$ si $\alpha \neq 1$).

No obstante, es posible expresar toda solución de un sistema no homogéneo compatible $A\boldsymbol{v} = \boldsymbol{b}$ como la suma de una solución **particular** de dicho sistema más una solución del sistema **homogéneo** (como se mencionó en el Cap. **1.**), es decir, del espacio nulo $N(A)$:

Teorema 4.4.1
Sea $A \in \mathbb{R}^{m\times n}$ y $\boldsymbol{b} \in \mathbb{R}^{m\times 1}$, con $\boldsymbol{v}_p \in \mathbb{R}^{n\times 1}$ una solución del sistema no homogéneo (asumido compatible)

$$A\boldsymbol{v} = \boldsymbol{b}$$

tal que $A\boldsymbol{v}_p = \boldsymbol{b}$. Entonces toda solución del sistema anterior es de la forma

$$\boldsymbol{v} = \boldsymbol{v}_p + \boldsymbol{v}_h$$

donde $\boldsymbol{v}_h \in N(A)$ es una solución del sistema homogéneo ($A\boldsymbol{v}_h = \boldsymbol{0}$).

Demostración. En primer lugar, es claro que si $\boldsymbol{v}_p$ es solución del sistema no homogéneo, también lo será $\boldsymbol{v}_p + \boldsymbol{v}_h$, con $\boldsymbol{v}_h$ cualquier solución del sistema homogéneo $A\boldsymbol{v}_h = \boldsymbol{0}$, o sea cualquier vector $\in N(A)$:

$$\begin{aligned} A(\boldsymbol{v}_p + \boldsymbol{v}_h) &= A\boldsymbol{v}_p + A\boldsymbol{v}_h \\ &= \boldsymbol{b} + \boldsymbol{0} = \boldsymbol{b} \end{aligned}$$

Y si $\boldsymbol{v}$ es cualquier otra solución del sistema no homogéneo ($A\boldsymbol{v} = \boldsymbol{b}$), entonces

$$A(\boldsymbol{v} - \boldsymbol{v}_p) = A\boldsymbol{v} - A\boldsymbol{v}_p = \boldsymbol{b} - \boldsymbol{b} = \boldsymbol{0}$$

por lo que $\boldsymbol{v} - \boldsymbol{v}_p$ es una solución del sistema homogéneo, es decir, $\boldsymbol{v} - \boldsymbol{v}_p = \boldsymbol{v}_h \in N(A)$. Por lo tanto, despejando $\boldsymbol{v}$ podemos expresar esta solución como

$$\boldsymbol{v} = \boldsymbol{v}_p + \boldsymbol{v}_h$$

Geométricamente, el conjunto de soluciones del sistema no homogéneo corresponde entonces a un "subespacio trasladado", tal como un plano o recta que no pasa por el origen.

Ejemplo 4.4.2: Sea A la matriz del ejemplo 4.4.1. Consideremos ahora el sistema no homogéneo $A\boldsymbol{v} = \boldsymbol{b}$, con $\boldsymbol{b} = \binom{1}{1}$. Aplicando la reducción de Gauss-Jordan obtenemos

$$(A \mid \boldsymbol{b}) = \left(\begin{array}{cccc|c} 1 & 1 & 0 & 1 & 1 \\ 2 & 3 & 1 & 2 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cccc|c} 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & -1 \end{array}\right) \longrightarrow \left(\begin{array}{cccc|c} 1 & 0 & -1 & 1 & 2 \\ 0 & 1 & 1 & 0 & -1 \end{array}\right)$$

Por lo tanto, el conjunto solución es

$$\boldsymbol{v} = \left\{ \begin{pmatrix} 2 + x_3 - x_4 \\ -1 - x_3 \\ x_3 \\ x_4 \end{pmatrix}, x_3, x_4 \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \ x_3, x_4 \in \mathbb{R} \right\}$$

Comparando con la solución del sistema homogéneo obtenida en el ejemplo **4.4.1**, vemos que toda solución del sistema no homogéneo es de la forma $\boldsymbol{v} = \boldsymbol{v}_p + \boldsymbol{v}_h$, con

$$\boldsymbol{v}_p = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \end{pmatrix}$$

una solución particular del sistema no homogéneo y

$$\boldsymbol{v}_h = x_3 \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} x_3 - x_4 \\ -x_3 \\ x_3 \\ x_4 \end{pmatrix}$$

una solución del sistema homogéneo (y por lo tanto $\in N(A)$). Si se resuelve primero el sistema no homogéneo, la solución del sistema homogéneo puede identificarse fácilmente como la parte de la solución dependiente de los parámetros libres (es decir, de las variables independientes).

Problemas 4.4
Hallar el espacio nulo de las siguientes matrices e interprételos geométricamente.
i) $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 3 \end{pmatrix}$ ii) $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 4 \end{pmatrix}$ iii) $A = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 0 & 2 \\ 3 & 0 & 3 \end{pmatrix}$

4.5. Espacio generado

Dado un conjunto de vectores $M = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ de un espacio vectorial V, se denomina ***espacio generado*** por M al conjunto de todas las combinaciones lineales de los elementos de M:

$$\text{gen}(M) = \{\boldsymbol{v} \in V : \boldsymbol{v} = \alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 + \ldots + \alpha_k \boldsymbol{v}_k, \ \alpha_i \in \mathbb{R}, i = 1, \ldots, k\}$$

Al espacio generado se lo indica como $\text{gen}(M)$ o también $\langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \rangle$.

Por ejemplo, el espacio nulo de la matriz A del ejemplo **4.4.1** es justamente el espacio generado por el conjunto de vectores $M = \left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$.

Teorema 4.5.1
El espacio generado $\text{gen}(M) = \langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \rangle$ es un subespacio de V

Demostración.

1. $\mathbf{0} \in \text{gen}(M)$, ya que se puede escribir como $\mathbf{0} = 0\,\boldsymbol{v}_1 + \ldots + 0\,\boldsymbol{v}_k$

2. Si $\boldsymbol{u} = \alpha_1\boldsymbol{v}_1 + \ldots + \alpha_k\boldsymbol{v}_k$ y $\boldsymbol{v} = \beta_1\boldsymbol{v}_1 + \ldots + \beta_k\boldsymbol{v}_k$ son dos vectores de $\text{gen}(M)$, por la propiedad distributiva (respecto de la suma de escalares) y la conmutatividad y asociatividad de la suma de vectores tenemos

$$\boldsymbol{u} + \boldsymbol{v} = (\alpha_1 + \beta_1)\boldsymbol{v}_1 + \ldots + (\alpha_k + \beta_k)\boldsymbol{v}_k \in \text{gen}(M)$$

por lo que $\text{gen}(M)$ es cerrado respecto a la suma de vectores.

3. Si $\boldsymbol{u} = \alpha_1\boldsymbol{v}_1 + \ldots + \alpha_k\boldsymbol{v}_k \in \text{gen}(M)$ y $\gamma \in \mathbb{R}$, por la propiedad distributiva (respecto de la suma de vectores) y la propiedad **7** tenemos

$$\gamma\boldsymbol{u} = \gamma(\alpha_1\boldsymbol{v}_1 + \ldots + \alpha_k\boldsymbol{v}_k) = (\gamma\alpha_1)\boldsymbol{v}_1 + \ldots + (\gamma\alpha_k)\boldsymbol{v}_k \in \text{gen}(M)$$

por lo que $\text{gen}(M)$ es cerrado con respecto al producto por un escalar.

Se concluye que $\text{gen}(M)$ es siempre un subespacio de V.

Ejemplos 4.5

1) Dado $M = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\} \subset \mathbb{R}^3$, el vector $\boldsymbol{v} = \begin{pmatrix} 3 \\ 5 \\ 2 \end{pmatrix} \in \text{gen}(M)$ ya que se puede escribir como

$$\begin{pmatrix} 3 \\ 5 \\ 2 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

El subespacio $\text{gen}(M) = \left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\rangle$ corresponde a un **plano de $\mathbb{R}^3$ que pasa por el origen de coordenadas** y cuya ecuación se puede determinar de la siguiente manera: Si $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \text{gen}(M) \Rightarrow \exists\ \alpha, \beta$ tal que

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha \\ \alpha + \beta \\ \beta \end{pmatrix}$$

Resolviendo el sistema se obtiene $\alpha = x$, $\beta = z$ y por lo tanto $y = \alpha + \beta = x + z$. Los vectores generados por M deben entonces satisfacer la ecuación $y = x + z$, es decir,

$$x - y + z = 0$$

que es la ecuación de un plano perpendicular al vector $(1,-1,1)$ que pasa por el origen.

Observación. Generalizando este ejemplo podemos afirmar que el espacio generado por dos vectores no nulos y no paralelos de $\mathbb{R}^3$ es siempre un plano que pasa por el origen.

2) Dado $M = \{1, t, t^2\} \subset P_2$, el polinomio $3t - 6t^2$ se puede escribir como combinación lineal de los elementos de M pues $3t - 6t^2 = 0\,1 + 3t + (-6)t^2$. De hecho todo polinomio de grado ≤ 2 puede escribirse como combinación lineal de los elementos de M.

3) Dado $M = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\} \subset \mathbb{R}^{2\times 2}$, toda matriz simétrica $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ $\in \text{gen}(M)$, pues

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix} = \frac{a+c}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{a-c}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Pero una matriz no simétrica $\in \mathbb{R}^{2\times 2}$ no puede ser generada por M (¡probar!).

4.6. Conjunto generador

Un conjunto $M = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ es un conjunto generador del espacio vectorial V si y sólo si todo vector de V puede escribirse como combinación lineal de $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$, o sea, si

$$V = \text{gen}(M) = \langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \rangle$$

Ejemplos 4.6

1) Sea $M = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} = \{\boldsymbol{e}_1, \boldsymbol{e}_2\} \subset \mathbb{R}^2$ ($\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$). Entonces

$$\text{gen}(M) = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle = \mathbb{R}^2$$

pues todo vector $\boldsymbol{v} = \begin{pmatrix} x \\ y \end{pmatrix}$ de $\mathbb{R}^2$ puede escribirse como combinación lineal de $\boldsymbol{e}_1$ y $\boldsymbol{e}_2$:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix} = x\boldsymbol{e}_1 + y\boldsymbol{e}_2$$

2) Sea $M = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} = \{\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}\} \subset \mathbb{R}^3$. Entonces

$$\text{gen}(M) = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle = \mathbb{R}^3$$

pues todo vector $\boldsymbol{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ de $\mathbb{R}^3$ se puede escribir como combinación lineal de $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = x\boldsymbol{e}_1 + y\boldsymbol{e}_2 + z\boldsymbol{e}_3$$

3) Sea $M = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\} \subset \mathbb{R}^{2\times 2}$. Entonces

$$\text{gen}(M) = \left\langle \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\rangle = \mathbb{R}^{2\times 2}$$

pues cualquier matriz $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}^{2\times 2}$ puede escribirse como

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

4) Sea $V = P_n$ y $M = \{1, t, t^2, \ldots, t^n\} \subset P_n$.
Todo polinomio $p(t) = a_0 + a_1 t + \ldots + a_n t^n \in P_n$ es una combinación lineal de estos monomios, por lo que

$$\text{gen}(M) = \langle 1, t, \ldots, t^n \rangle = P_n$$

5) Sea $M = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\} \subset \mathbb{R}^2$. Dado un vector cualquiera $\boldsymbol{v} = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$, ¿es posible escribirlo como combinación lineal de los vectores de M? Para ello debemos ver si existen escalares α, β tales que

$$\begin{pmatrix} x \\ y \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Se obtiene así el sistema lineal $\begin{cases} \alpha + \beta & = & x \\ \beta & = & y \end{cases}$ cuya solución es $\beta = y$, $\alpha = x - y$.
Por lo tanto,

$$\begin{pmatrix} x \\ y \end{pmatrix} = (x - y) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

lo que implica que este conjunto también genera $\mathbb{R}^2$:

$$\left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle = \mathbb{R}^2$$

Geométricamente, esto se debe a que los dos vectores de M no son colineales, como veremos en breve.

6) Sea $M' = \left\{ \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} \right\} \subset \mathbb{R}^3$.

Todo vector $\boldsymbol{v} \in \mathbb{R}^3$ puede también escribirse como

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + 0 \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$$

por lo cual también $\text{gen}(M') = \left\langle \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} \right\rangle = \mathbb{R}^3$. Podemos observar que el último vector de M' es "redundante" ya que no es necesario para obtener al vector $\boldsymbol{v}$. En realidad, puede quitarse uno cualquiera de estos cuatro vectores sin afectar la generación de todo $\mathbb{R}^3$ (¡probar!)

7) Sea

$$M = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 3 \\ 2 \end{pmatrix} \right\}$$

Dado un vector cualquiera $\mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$, ¿podemos escribirlo como combinación lineal de los vectores de M? Es decir, ¿genera M a $\mathbb{R}^3$? Para ello debemos ver si existen escalares α, β y δ tales que

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + \delta \begin{pmatrix} 4 \\ 3 \\ 2 \end{pmatrix}$$

Resolviendo este sistema por eliminación gaussiana,

$$\begin{aligned} (A \mid \mathbf{v}) \;\; &= \left(\begin{array}{ccc|c} 1 & 2 & 4 & x \\ 1 & 1 & 3 & y \\ 1 & 0 & 2 & z \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & 4 & x \\ 0 & -1 & -1 & y-x \\ 0 & -2 & -2 & z-x \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & 4 & x \\ 0 & 1 & 1 & x-y \\ 0 & -2 & -2 & z-x \end{array}\right) \\ &\to \left(\begin{array}{ccc|c} 1 & 2 & 4 & x \\ 0 & 1 & 1 & x-y \\ 0 & 0 & 0 & x-2y+z \end{array}\right) \end{aligned}$$

vemos que si $x - 2y + z \neq 0$ el sistema es incompatible. Los vectores $\mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ para los cuales $x - 2y + z \neq 0$ no pueden ser generados por el conjunto M.

Por otro lado, los vectores $\boldsymbol{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ que satisfacen $x - 2y + z = 0$ sí pueden ser generados por M, teniendo el sistema infinitas soluciones.

Es decir que M no genera $\mathbb{R}^3$, sino un subespacio propio S que es un plano que pasa por el origen, definido por la ecuación

$$x - 2y + z = 0$$

Sólo aquellos vectores que están en este plano son generados por M: $S = \text{gen}(M)$.

Problemas 4.6 Indique si los siguientes conjuntos generan $\mathbb{R}^3$:

a) $M = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$ b) $M = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \right\}$

4.6.1. Conjunto generador minimal

Consideremos nuevamente el ejemplo **7)** anterior. Para generar un plano sólo se necesitan dos vectores no nulos y no paralelos pertenecientes al plano, por lo cual, en realidad, no se necesitan los tres vectores de M para generar S. Uno de ellos es innecesario o redundante, perteneciendo al plano ya generado por los otros dos.

Para decidir si un conjunto generador M de un espacio vectorial V constituye un conjunto generador minimal, es decir, sin elementos redundantes, debemos analizar si los vectores de M dependen linealmente unos de los otros.

Volviendo al ejemplo **7)** anterior, podemos observar que $\begin{pmatrix} 4 \\ 3 \\ 2 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$

Es decir, si $\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $\boldsymbol{v}_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$ y $\boldsymbol{v}_3 = \begin{pmatrix} 4 \\ 3 \\ 2 \end{pmatrix}$, tenemos $\boldsymbol{v}_3 = 2\boldsymbol{v}_1 + \boldsymbol{v}_2$. Geométricamente, $\boldsymbol{v}_3$ está en el plano generado por $\boldsymbol{v}_1$ y $\boldsymbol{v}_2$.

Entonces $S = \text{gen}(M) = \langle \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3 \rangle = \langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle$, ya que cualquier combinación lineal de $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ puede ser reducida a una combinación lineal de $\boldsymbol{v}_1$ y $\boldsymbol{v}_2$:

$$\alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \alpha_3\boldsymbol{v}_3 = \alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \alpha_3(2\boldsymbol{v}_1 + \boldsymbol{v}_2) = (\alpha_1 + 2\alpha_3)\boldsymbol{v}_1 + (\alpha_2 + 1)\boldsymbol{v}_2 = \beta_1\boldsymbol{v}_1 + \beta_2\boldsymbol{v}_2$$

Podemos reescribir la dependencia de $\boldsymbol{v}_3$ con respecto de $\boldsymbol{v}_1$ y $\boldsymbol{v}_2$ como

$$2\boldsymbol{v}_1 + \boldsymbol{v}_2 - \boldsymbol{v}_3 = 0$$

Como ninguno de los tres coeficientes de $\boldsymbol{v}_1, \boldsymbol{v}_2$ y $\boldsymbol{v}_3$ es nulo se puede despejar a cualquiera de los vectores en función de los dos restantes. Por lo tanto, tenemos también

$$S = \text{gen}(M) = \langle \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3 \rangle = \langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle = \langle \boldsymbol{v}_2, \boldsymbol{v}_3 \rangle = \langle \boldsymbol{v}_1, \boldsymbol{v}_3 \rangle$$

En este ejemplo, cualquiera de los tres vectores de M puede ser considerado redundante, ya que puede ser expresado como combinación lineal de los dos restantes y con sólo dos vectores se puede generar S. Vemos también que ningun vector es proporcional a otro. Geométricamente, se trata de tres vectores no paralelos, pero situados **en un mismo plano**. Por lo tanto, $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ no es un conjunto generador minimal de S, pero $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$, $\{\boldsymbol{v}_1, \boldsymbol{v}_3\}$ y $\{\boldsymbol{v}_2, \boldsymbol{v}_3\}$ son conjuntos generadores minimales del plano S.

El siguiente teorema generalizan lo visto en el ejemplo anterior:

Teorema 4.6.1
Si $V = \langle \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k \rangle$ $(k \geq 2)$ y alguno de los $\boldsymbol{v}_i$ puede ser escrito como una combinación lineal de los restantes $(k-1)$ vectores, entonces estos $(k-1)$ vectores ya generan V.

Demostración.
Supongamos que $\boldsymbol{v}_k$ puede ser escrito como combinación lineal de $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{k-1}$:

$$\boldsymbol{v}_k = \beta_1 \boldsymbol{v}_1 + \ldots + \beta_{k-1} \boldsymbol{v}_{k-1}$$

entonces toda combinación lineal de estos k vectores puede ser reducida a una combinacón lineal de los primeros $k-1$ vectores:

$$\begin{aligned}\alpha_1 \boldsymbol{v}_1 + \ldots + \alpha_{k-1} \boldsymbol{v}_{k-1} + \alpha_k \boldsymbol{v}_k &= \alpha_1 \boldsymbol{v}_1 + \ldots + \alpha_{k-1} \boldsymbol{v}_{k-1} + \alpha_k (\beta_1 \boldsymbol{v}_1 + \ldots + \beta_{k-1} \boldsymbol{v}_{k-1}) \\ &= (\alpha_1 + \alpha_k \beta_1) \boldsymbol{v}_1 + \ldots + (\alpha_{k-1} + \alpha_k \beta_{k-1}) \boldsymbol{v}_{k-1}\end{aligned}$$

Esto implica

$$\langle \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k \rangle = \langle \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{k-1} \rangle$$

Todo $\boldsymbol{v} \in V$ puede pues escribirse como combinación lineal de los primeros $k-1$ vectores. Para que $M = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ sea un conjunto generador minimal del espacio que genera es necesario que ningún vector sea combinación lineal de los restantes, es decir, que los k vectores sean **linealmente independientes**, como discutiremos a continuación.

4.7. Independencia lineal

Sea V un espacio vectorial. El conjunto no vacío $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\} \subset V$ es ***linealmente independiente*** si la ecuación

$$\alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 + \ldots + \alpha_k \boldsymbol{v}_k = \boldsymbol{0}$$

implica necesariamente que todos los escalares α_k sean nulos:

$$\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$$

Por el contrario, el conjunto $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\} \subset V$ es ***linealmente dependiente*** si existen escalares $\alpha_1, \alpha_2, \ldots, \alpha_k$ no todos nulos tales que

$$\alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 + \ldots + \alpha_k \boldsymbol{v}_k = \boldsymbol{0}$$

El siguiente teorema muestra el significado de estas definiciones:

Teorema 4.7.1
Dados k vectores $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ $(k \geq 2)$, al menos uno de estos vectores es combinación lineal de los restantes $k-1$ vectores si y sólo si los vectores son linealmente dependientes.

Demostración.
$\Rightarrow$: Supongamos que uno de los k vectores es combinación lineal de los restantes, por ej.

$\boldsymbol{v}_k = \beta_1 \boldsymbol{v}_1 + \ldots + \beta_{k-1}\boldsymbol{v}_{k-1}$. Entonces, restando $\boldsymbol{v}_k$ en ambos miembros,

$$\beta_1 \boldsymbol{v}_1 + \ldots + \beta_{k-1}\boldsymbol{v}_{k-1} - \boldsymbol{v}_k = \mathbf{0}$$

por lo que existe una combinación lineal de los k vectores con coeficientes no todos nulos que es nula ($\alpha_i = \beta_i$ si $i \leq k-1$, $\alpha_k = -1$).
$\Leftarrow$: Si $\alpha_1 \boldsymbol{v}_1 + \ldots + \alpha_k \boldsymbol{v}_k = \mathbf{0}$ y los α_i no son todos nulos, suponiendo por ejemplo $\alpha_k \neq 0$ podemos despejar $\boldsymbol{v}_k$ en términos de los restantes vectores:

$$\boldsymbol{v}_k = -(\alpha_1 \boldsymbol{v}_1 + \ldots + \alpha_{k-1}\boldsymbol{v}_{k-1})/\alpha_k = (-\tfrac{\alpha_1}{\alpha_k})\boldsymbol{v}_1 + \ldots + (-\tfrac{\alpha_{k-1}}{\alpha_k})\boldsymbol{v}_{k-1}$$

lo que muestra que $\boldsymbol{v}_k$ es una combinación lineal de los restantes $k-1$ vectores.

Los vectores $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ son entonces **linealmente dependientes** si existe una combinación lineal de ellos con coeficientes no todos nulos que es nula. En este caso al menos uno de los k vectores $\boldsymbol{v}_i$ pertenece al espacio generado por los restantes.

Por el contrario, los vectores $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ son **linealmente independientes** si la única forma de lograr una combinación lineal nula es que todos los escalares α_i sean nulos. En este caso, ninguno de los k vectores $\boldsymbol{v}_i$ puede ser escrito como combinación lineal de los restantes, es decir, ninguno pertenece al espacio generado por los restantes.

Si $k = 1$, las definiciones anteriores implican:

$$\begin{cases} \boldsymbol{v} = \mathbf{0} \;\Leftrightarrow\; \{\boldsymbol{v}\} \text{ linealmente dependiente} \\ \boldsymbol{v} \neq \mathbf{0} \;\Leftrightarrow\; \{\boldsymbol{v}\} \text{ linealmente independiente} \end{cases}$$

ya que si $\boldsymbol{v} = \mathbf{0}$, $\alpha\boldsymbol{v} = \mathbf{0}\ \forall\ \alpha$ mientras que si $\boldsymbol{v} \neq \mathbf{0}$, $\alpha\boldsymbol{v} = \mathbf{0}$ implica $\alpha = 0$ (Teorema 4.2.1).

Si $k = 2$, tenemos

$$\begin{cases} \boldsymbol{v}_1 \text{ y } \boldsymbol{v}_2 \text{ proporcionales (colineales)} \;\Leftrightarrow\; \{\boldsymbol{v}_1, \boldsymbol{v}_2\} \text{ linealmente dependiente} \\ \boldsymbol{v}_1 \text{ y } \boldsymbol{v}_2 \text{ no proporcionales (y no nulos)} \;\Leftrightarrow\; \{\boldsymbol{v}_1, \boldsymbol{v}_2\} \text{ linealmente independiente} \end{cases}$$

ya que si son proporcionales, por ej. $\boldsymbol{v}_2 = \alpha\boldsymbol{v}_1$, son linealmente dependientes ($\boldsymbol{v}_2 - \alpha\boldsymbol{v}_1 = \mathbf{0}$), mientras que si

$$\alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 = \mathbf{0}$$

con α_1 o α_2 (o ambos) no nulos (vectores linealmente dependientes) entonces

$$\boldsymbol{v}_2 = -\frac{\alpha_1}{\alpha_2}\boldsymbol{v}_1 \;\; (\alpha_2 \neq 0) \;\text{ o }\; \boldsymbol{v}_1 = -\frac{\alpha_2}{\alpha_1}\boldsymbol{v}_2 \;\; (\alpha_1 \neq 0)$$

que implica que son necesariamente proporcionales, es decir, uno es un multiplo escalar del otro. Esto incluye el caso en que $\boldsymbol{v}_1$ o $\boldsymbol{v}_2$ (o ambos) son nulos (si $\boldsymbol{v}_2 = \mathbf{0} \Rightarrow \boldsymbol{v}_2 = 0\boldsymbol{v}_1$).

Geométricamente, en $V = \mathbb{R}^n$ esto significa que los vectores $\boldsymbol{v}_1$ y $\boldsymbol{v}_2$ serán linealmente dependientes sólo si son **colineales** ($\boldsymbol{v}_2 \propto \boldsymbol{v}_1$ o $\boldsymbol{v}_1 \propto \boldsymbol{v}_2$), es decir, si pertenecen ambos a una misma recta que pasa por el origen, incluyendo el caso en que uno o ambos son

nulos. Por el contrario, los vectores $\boldsymbol{v}_1$ y $\boldsymbol{v}_2$ serán linealmente independientes si no son colineales. Nótese que esto no implica que sean ortogonales.

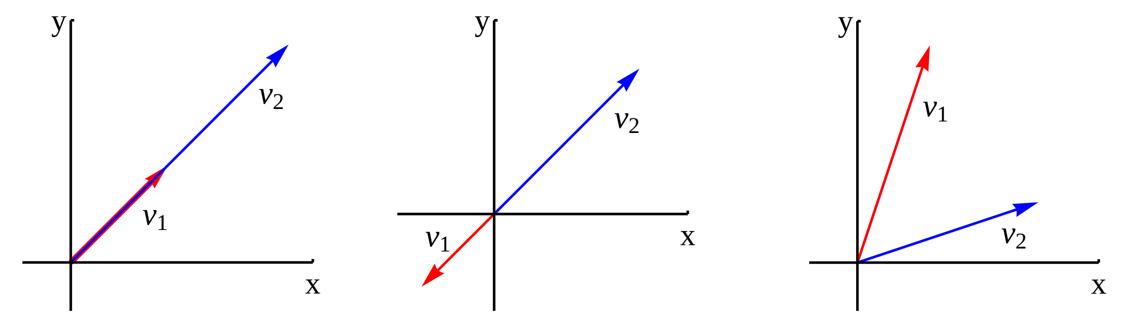


Figura 4.5: vectores linealmente dependientes vectores linealmente independientes

En $\mathbb{R}^3$, dos vectores $\boldsymbol{v}_1$ y $\boldsymbol{v}_2$ que sean linealmente independientes generan un plano, pues son no nulos y no colineales. Si un tercer vector $\boldsymbol{v}_3$ pertenece a ese plano, de forma que $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ son coplanares, el conjunto $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ será **linealmente dependiente**, ya que $\boldsymbol{v}_3$ podrá escribirse como combinación lineal de $\boldsymbol{v}_1$ y $\boldsymbol{v}_2$. Por otro lado, si los tres vectores no son coplanares, el conjunto $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ será **linealmente independiente** y generará todo $\mathbb{R}^3$, como veremos luego.

Mencionemos finalmente que si uno de los vectores del conjunto $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ es nulo entonces $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ es **linealmente dependiente**: Suponiendo, por ej., $\boldsymbol{v}_k = \mathbf{0}$, tenemos $0\boldsymbol{v}_1 + \ldots + 0\boldsymbol{v}_{k-1} + 1\boldsymbol{v}_k = \mathbf{0} + 1\mathbf{0} = \mathbf{0}$, existiendo entonces una combinación lineal nula de los vectores con coeficientes no todos nulos.

Ejemplos 4.7.1

1) Sea $M = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$. Planteando la combinación lineal nula

$$\alpha_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

obtenemos el sistema $\left\{ \begin{aligned} \alpha_1 + \alpha_2 + \alpha_3 &= 0 \\ \alpha_2 + \alpha_3 &= 0 \\ \alpha_3 &= 0 \end{aligned} \right.$, es decir $\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right)$

que tiene como única solución $\alpha_1 = \alpha_2 = \alpha_3 = 0$. El conjunto M es entonces linealmente **independiente**. Este resultado es también obvio a simple vista: Por la forma de los vectores, es claro que ninguno puede escribirse como combinación lineal de los otros dos.

Nótese también que la matriz $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ formada por los tres vectores es **no singular** ($\det A = 1$).

2) Sea $M = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 3 \\ 2 \end{pmatrix} \right\}$. Planteando la combinación lineal nula

$$\alpha_1\begin{pmatrix}1\\1\\1\end{pmatrix}+\alpha_2\begin{pmatrix}2\\1\\0\end{pmatrix}+\alpha_3\begin{pmatrix}4\\3\\2\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

obtenemos el sistema $\begin{cases}\alpha_1+2\alpha_2+4\alpha_3=0\\ \alpha_1+\alpha_2+3\alpha_3=0\\ \alpha_1+2\alpha_3=0\end{cases}$, es decir

$$\left(\begin{array}{ccc|c}1&2&4&0\\1&1&3&0\\1&0&2&0\end{array}\right)\longrightarrow\left(\begin{array}{ccc|c}1&2&4&0\\0&1&1&0\\0&0&0&0\end{array}\right)\longrightarrow\left(\begin{array}{ccc|c}1&0&2&0\\0&1&1&0\\0&0&0&0\end{array}\right)$$

que tiene infinitas soluciones: $\alpha_1=-2\alpha_3$, $\alpha_2=-\alpha_3$, con α_3 libre. Esto implica

$$\alpha_3\left[-2\begin{pmatrix}1\\1\\1\end{pmatrix}-\begin{pmatrix}2\\1\\0\end{pmatrix}+\begin{pmatrix}4\\3\\2\end{pmatrix}\right]=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

$\forall\ \alpha_3$. El conjunto M es entonces linealmente **dependiente**.

Nótese que la matriz $A=\begin{pmatrix}1&2&4\\1&1&3\\1&0&2\end{pmatrix}$ formada por los tres vectores es **singular** ($\det A=0$).

3) Sea $M=\left\{\boldsymbol{u}=\begin{pmatrix}2\\-1\\3\end{pmatrix},\ \boldsymbol{v}=\begin{pmatrix}-4\\2\\-6\end{pmatrix}\right\}$. Claramente se observa que $\boldsymbol{v}=-2\boldsymbol{u}$, por lo cual estos dos vectores son linealmente **dependientes**. Como verificación, la ecuación $\alpha\boldsymbol{u}+\beta\boldsymbol{v}=\boldsymbol{0}$ conduce al sistema $\begin{cases}2\alpha-4\beta&=0\\-\alpha+2\beta&=0\\3\alpha-6\beta&=0\end{cases}$, o sea,

$$\left(\begin{array}{cc|c}2&-4&0\\-1&2&0\\3&-6&0\end{array}\right)\longrightarrow\left(\begin{array}{cc|c}2&-4&0\\0&0&0\\0&0&0\end{array}\right)\longrightarrow\left(\begin{array}{cc|c}1&-2&0\\0&0&0\\0&0&0\end{array}\right)$$

que implica $\alpha=2\beta$ con β libre, es decir, $\beta(2\boldsymbol{u}+\boldsymbol{v})=\boldsymbol{0}\ \forall\ \beta$, por lo que $\boldsymbol{v}=-2\boldsymbol{u}$.

4) Sea $M=\{1+t, t+t^2, 2+3t+2t^2\}\subset P_2$. Para ver si son polinomios linealmente independientes, consideramos la ecuación

$$\alpha_1(1+t)+\alpha_2(t+t^2)+\alpha_3(2+3t+t^2)=0$$

es decir, $\alpha_1+2\alpha_3+(\alpha_1+\alpha_2+3\alpha_3)t+(\alpha_2+\alpha_3)t^2=0+0t+0t^2$, donde la igualdad debe valer $\forall\ t$. Esto conduce al sistema $\begin{cases}\alpha_1+2\alpha_3&=&0\\\alpha_1+\alpha_2+3\alpha_3&=&0\\\alpha_2+\alpha_3&=&0\end{cases}$ que es compatible indeterminado, siendo el conjunto solución $\alpha_2=-\alpha_3$, $\alpha_1=-2\alpha_3$, con α_3 libre. Esto implica que son linealmente dependientes, con $-2(1+t)-(t+t^2)+(2+3t+t^2)=0$.

Nótese que la matriz $A=\begin{pmatrix}1&0&2\\1&1&3\\0&1&1\end{pmatrix}$ es singular ($\det(A)=0$).

5) La indepencia lineal de funciones es un concepto importante en la teoría de ecuaciones diferenciales lineales, como veremos en el 2° módulo.
Consideremos, por ejemplo, las funciones $f_1(t) = \cos(t)$, $f_2(t) = \mathrm{sen}(t)$ incluidas en $V = C(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R},\ f \text{ continua}\}$. Es obvio que son **linealmente independientes**, pues no son proporcionales: $\mathrm{sen}(t) \neq \alpha \cos(t)\ \forall\ t \in \mathbb{R}$, es decir, el cociente $\frac{\mathrm{sen}(t)}{\cos(t)} = \tan(t)$ es la función tangente, que no es una función constante.
Esto se puede también probar formalmente planteando la ecuación

$$\alpha_1 \cos(t) + \alpha_2 \,\mathrm{sen}(t) = 0$$

que debe ser válida $\forall\ t \in \mathbb{R}$. Considerando por ej. $t = 0$ y $t = \pi/2$, obtenemos el sistema $\left\{ \begin{array}{rcl} \alpha_1 + 0\,\alpha_2 & = & 0 \\ 0\,\alpha_1 + \alpha_2 & = & 0 \end{array} \right.$, que conduce a la solución única $\alpha_1 = \alpha_2 = 0$.

6) Consideremos ahora el conjunto de funciones $M = \{\cos^2(t), \mathrm{sen}^2(t), 1\} \subset C(\mathbb{R})$ (1 denota la función constante $f(t) = 1\ \forall\ t$). Dado que $\cos^2(t) + \mathrm{sen}^2(t) = 1$, tenemos

$$\cos^2(t) +\ \mathrm{sen}^2(t) - 1 = 0$$

$\forall\ t \in \mathbb{R}$, por lo que el conjunto es **linealmente dependiente**. Cualquiera de estas tres funciones puede escribirse como combinación lineal de las otras dos.

Propiedades fundamentales

En los ejemplos anteriores vimos que en el caso de tres vectores en $\mathbb{R}^3$, si la matriz A formada por las coordenadas (coeficientes) de los tres vectores es no singular el conjunto es linealmente independiente, mientras que si A es singular el conjunto es linealmente dependiente. Este resultado se generaliza a $\mathbb{R}^n$:

Teorema 4.7.2:

Sean $\boldsymbol{v}_1 = \begin{pmatrix} v_{11} \\ \vdots \\ v_{n1} \end{pmatrix}, \ldots, \boldsymbol{v}_n = \begin{pmatrix} v_{1n} \\ \vdots \\ v_{nn} \end{pmatrix}$ n vectores de $\mathbb{R}^n$.

El conjunto $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ es **linealmente independiente** si y sólo si la matriz de $n \times n$

$$A = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) = \begin{pmatrix} v_{11} & \ldots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{n1} & \ldots & v_{nn} \end{pmatrix}$$

es no singular, es decir, $\det A \neq 0$. Esto también implica que el conjunto $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ es **linealmente dependiente** si y sólo si la matriz A es singular, es decir, $\det A = 0$.

Demostración.

La combinación lineal nula $\alpha_1 \boldsymbol{v}_1 + \ldots + \alpha_n \boldsymbol{v}_n = \mathbf{0}$ conduce al sistema homogéneo

$$\alpha_1 \begin{pmatrix} v_{11} \\ \vdots \\ v_{n1} \end{pmatrix} + \ldots + \alpha_n \begin{pmatrix} v_{1n} \\ \vdots \\ v_{nn} \end{pmatrix} = \begin{pmatrix} v_{11} & \ldots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{n1} & \ldots & v_{nn} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

es decir, $A\boldsymbol{\alpha} = \mathbf{0}$ con A la matriz anterior y $\boldsymbol{\alpha}$ un vector columna de $n \times 1$ de componentes α_i (incógnitas). Si A es no singular ($\det A \neq 0$), este sistema tendrá como única solución la solución trivial $\alpha_1 = \ldots = \alpha_n = 0$, en cuyo caso el conjunto $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ será linealmente independiente. Por el contrario, si A es singular ($\det A = 0$), el sistema será compatible indeterminado, con infintas soluciones no triviales para $\boldsymbol{\alpha}$, en cuyo caso el conjunto $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ será linealmente dependiente.

Geométricamente, y recordando que $|\det A|$ es el "volumen" del paralelepípedo formado por las columnas (o filas) de A (es decir, el área en $\mathbb{R}^2$, el volumen en $\mathbb{R}^3$, etc.), vemos que un conjunto linealmente independiente de vectores forma un paralelepípedo de **volumen no nulo**, mientras que un conjunto linealmente dependiente genera un **volumen nulo** por ser los vectores "coplanares".

Así, una forma de determinar si n vectores de $\mathbb{R}^n$ son linealmente independientes es construir la matriz A de $n \times n$ colocando en cualquier orden a los vectores como columnas de A y luego calcular $\det A$ (dado que $\det A = \det A^T$, se obtiene el mismo resultado si se los coloca por filas). Resumiendo,

- $\det A \neq 0 \Leftrightarrow \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ linealmente independiente
- $\det A = 0 \Leftrightarrow \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ linealmente dependiente

Ejemplo. El conjunto de vectores $M = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \right\}$ es linealmente independiente pues $\det A = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 0 & 3 & 2 \end{vmatrix} = 4 - 9 = -5 \neq 0$.

Teorema 4.7.3
Todo conjunto de $m > n$ vectores en $\mathbb{R}^n$ es **linealmente dependiente**.

Demostración.
Sea $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ un conjunto de m vectores en $\mathbb{R}^n$, con $m > n$. La combinación lineal nula $\alpha_1 \boldsymbol{v}_1 + \ldots + \alpha_m \boldsymbol{v}_m = \mathbf{0}$ conduce al sistema lineal homogéneo

$$\alpha_1 \begin{pmatrix} v_{11} \\ \vdots \\ v_{n1} \end{pmatrix} + \ldots + \alpha_m \begin{pmatrix} v_{1m} \\ \vdots \\ v_{nm} \end{pmatrix} = \begin{pmatrix} v_{11} & \ldots & v_{1m} \\ \vdots & \ddots & \vdots \\ v_{n1} & \ldots & v_{nm} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

que es un sistema de n ecuaciones homogéneas con $m > n$ incógnitas $\alpha_1, \ldots, \alpha_m$. Como se vió en el capítulo de sistemas lineales, tales sistemas son siempre **compatibles indeterminados**, teniendo por tanto soluciones no triviales. Esto implica que el conjunto $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ será linealmente dependiente. Como consecuencia,

Todo conjunto linealmente independiente de vectores de $\mathbb{R}^n$ contiene a lo sumo n vectores.

Por ejemplo, en $\mathbb{R}^2$ podemos tener a lo sumo 2 vectores linealmente independientes y en $\mathbb{R}^3$ a lo sumo 3 vectores linealmente independientes.

Observación. Estos resultados son válidos en todo espacio vectorial de dimensión n, como veremos en la próxima sección.

Ejemplo: El conjunto $M = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ es linealmente dependiente pues se trata de 4 vectores de $\mathbb{R}^3$.

Unicidad. Un resultado importante para un conjunto M de k vectores linealmente independiente es que todo vector $\boldsymbol{v}$ perteneciente al espacio generado por M se puede escribir de manera **única** como combinación lineal de los vectores de M. Esto generaliza la unicidad de los escalares α_i vista en el teorema **4.7.2** para n vectores de $\mathbb{R}^n$ a un conjunto linealmente independiente arbitrario de vectores.

Teorema 4.7.4 Sea $M = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ un conjunto de k vectores de un espacio vectorial V y sea

$$\boldsymbol{v} = \alpha_1 \boldsymbol{v}_1 + \ldots + \alpha_n \boldsymbol{v}_k$$

un vector $\in \text{gen}(M) = \langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \rangle$. Los escalares $\alpha_1, \ldots, \alpha_k$ que determinan $\boldsymbol{v}$ son **únicos** si y sólo si $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ es **linealmente independiente.**

Demostración.
$\Leftarrow$) Supongamos $M = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ linealmente independiente y $\boldsymbol{v} \in \text{gen}(M)$. Si $\boldsymbol{v}$ puede escribirse de dos formas diferentes,

$$\begin{aligned} \boldsymbol{v} &= \alpha_1 \boldsymbol{v}_1 + \ldots + \alpha_k \boldsymbol{v}_k \\ &= \beta_1 \boldsymbol{v}_1 + \ldots + \beta_k \boldsymbol{v}_k \end{aligned}$$

restando obtenemos

$$\mathbf{0} = (\alpha_1 - \beta_1)\boldsymbol{v}_k + \ldots + (\alpha_k - \beta_k)\boldsymbol{v}_k$$

Pero como $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ es linealmente independiente $\Rightarrow \alpha_1 - \beta_1 = 0, \ldots, \alpha_k - \beta_k = 0$, por lo que $\alpha_i = \beta_i \ \forall \ i = 1, \ldots, k$. Es decir, los coeficientes son únicos.

$\Rightarrow$) Supongamos que todo vector $\boldsymbol{v} \in \text{gen}(M)$ puede escribirse de manera única como combinación lineal de $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$. Esto incluye al vector nulo $\mathbf{0}$, que puede escribirse como

$$\mathbf{0} = 0\boldsymbol{v}_1 + \ldots + 0\boldsymbol{v}_k$$

Como por hipótesis la representación debe ser única, no existe entonces una combinación lineal de los $\boldsymbol{v}_i$ con coeficientes no todos nulos que sea el vector nulo. Esto implica que los k vectores $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ son linealmente independientes.

El teorema también implica que si M es linealmente dependiente, existen multiples representaciones de un vector $\boldsymbol{v} \in \text{gen}(M)$ como combinación lineal de los $\boldsymbol{v}_i$.

Problemas 4.7

1) Analizar si los siguientes conjuntos son linealmente dependientes o independientes. En el caso dependiente mostrar la dependencia lineal.

a) i) $\left\{\begin{pmatrix}2\\1\end{pmatrix},\begin{pmatrix}1\\2\end{pmatrix}\right\}$ ii) $\left\{\begin{pmatrix}2\\1\end{pmatrix},\begin{pmatrix}1\\2\end{pmatrix},\begin{pmatrix}1\\0\end{pmatrix}\right\}$, ii) $\left\{\begin{pmatrix}1\\1\end{pmatrix},\begin{pmatrix}0\\0\end{pmatrix}\right\}$,

b) i) $\left\{\begin{pmatrix}1\\0\\1\end{pmatrix},\begin{pmatrix}1\\2\\1\end{pmatrix}\right\}$, ii) $\left\{\begin{pmatrix}1\\0\\1\end{pmatrix},\begin{pmatrix}1\\2\\1\end{pmatrix},\begin{pmatrix}1\\1\\1\end{pmatrix}\right\}$, iii) $\left\{\begin{pmatrix}1\\0\\1\end{pmatrix},\begin{pmatrix}1\\2\\1\end{pmatrix},\begin{pmatrix}0\\1\\1\end{pmatrix}\right\}$

c) i) $\left\{\begin{pmatrix}1&0\\0&1\end{pmatrix},\begin{pmatrix}1&0\\0&-1\end{pmatrix},\begin{pmatrix}0&1\\1&0\end{pmatrix}\right\}$, ii) $\left\{\begin{pmatrix}1&1\\1&1\end{pmatrix},\begin{pmatrix}1&-1\\-1&1\end{pmatrix},\begin{pmatrix}1&3\\3&1\end{pmatrix}\right\}$

d) i) $\{(1,0,1),(0,2,1),(0,0,3)\}$, ii) $\{(1,2,3,4),(0,2,3,4),(1,-2,-3,-4)\}\}$

2) a) Muestre que el conjunto $\{\boldsymbol{e}_1,\boldsymbol{e}_2\}=\left\{\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right\}\subset\mathbb{R}^2$ es linealmente independiente.

b) Extienda el resultado anterior a $\mathbb{R}^n$: Muestre que

$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\ldots,\boldsymbol{e}_n\}=\left\{\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix},\ldots,\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}\right\}\subset\mathbb{R}^n$ es linealmente independiente.

3) Mostrar que el vector $\begin{pmatrix}1\\4\\1\end{pmatrix}$ pertenece al espacio generado por el conjunto $\left\{\begin{pmatrix}1\\0\\1\end{pmatrix},\begin{pmatrix}1\\2\\1\end{pmatrix}\right\}$, encontrando los coeficientes α_1 y α_2 de la combinación lineal. ¿ Son únicos?

4) ¿ Para qué valores de c es el conjunto $M=\left\{\begin{pmatrix}1\\0\\1\end{pmatrix},\begin{pmatrix}1\\2\\0\end{pmatrix},\begin{pmatrix}c\\1\\0\end{pmatrix}\right\}$ linealmente dependiente?

5) Encuentre un conjunto de tres vectores linealmente independiente de $\mathbb{R}^3$ que contenga a los vectores $(1,2,3)$ y $(1,1,1)$.

6) En $V=C(0,\infty)$, analizar si son linealmente independientes los siguientes conjuntos de funciones: a) $\{\ln t, e^t\}$, b) $\{\ln(t^3),\ln(t)\}$, c) $\{\cos(2t),\text{sen}^2(t),1\}$

7) Si el conjunto de vectores $M=\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\}\subset V$ es linealmente independiente,
i) Muestre que el conjunto $\{\boldsymbol{u},\boldsymbol{u}+2\boldsymbol{v},\boldsymbol{u}+2\boldsymbol{v}+3\boldsymbol{w}\}$ es linealmente independiente.
ii) Muestre que los subconjuntos propios $\{\boldsymbol{u},\boldsymbol{v}\}$, $\{\boldsymbol{u},\boldsymbol{w}\}$, $\{\boldsymbol{v},\boldsymbol{w}\}$, $\{\boldsymbol{u}\}$, $\{\boldsymbol{v}\}$, $\{\boldsymbol{w}\}$, son todos linealmente independientes.
iii) ¿Es válida la recíproca? Si todos los subconjuntos propios anteriores son linealmente independientes, ¿es $\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\}$ necesariamente linealmente independiente?

8) Muestre que si $\{\boldsymbol{v}_1,\ldots,\boldsymbol{v}_k\}\subset V$ es linealmente dependiente, entonces $\{\boldsymbol{v}_1,\ldots,\boldsymbol{v}_k,\boldsymbol{v}_{k+1}\}$ es linealmente dependiente $\forall\ \boldsymbol{v}_{k+1}\in V$.

9) Mostrar que las filas no nulas de una matriz en forma escalonada reducida forman siempre un conjunto linealmente independiente de vectores.

10) Recordemos que el producto escalar entre dos vectores $\boldsymbol{u} = (u_1, \ldots, u_n)$, $\boldsymbol{v} = (v_1, \ldots, v_n)$ de $\mathbb{R}^n$ se define como $\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + \ldots + u_n v_n$.
a) Muestre que si $\boldsymbol{u}$ y $\boldsymbol{v}$ son no nulos y ortogonales ($\boldsymbol{u} \cdot \boldsymbol{v} = 0$) entonces son linealmente independientes (considere $n \geq 2$).
b) Si $\boldsymbol{u}$ y $\boldsymbol{v}$ son linealmente independientes, ¿son necesariamente ortogonales?
c) Generalizar a) al caso de $m \leq n$ vectores no nulos y mutuamente ortogonales. ¿Podría ser $m > n$?

11) Muestre que si dos funciones $f, g \in C^1(\mathbb{R})$ (el espacio de funciones $f : \mathbb{R} \to \mathbb{R}$ derivables) son linealmente dependientes, entonces el determinante (llamado *wronskiano*) $W(t) = \begin{vmatrix} f(t) & g(t) \\ f'(t) & g'(t) \end{vmatrix}$ es 0 $\forall\, t \in \mathbb{R}$. ¿Es válida la propiedad recíproca?

4.8. Bases y dimensión de un espacio vectorial

Un espacio vectorial V se dice que es **finitamente generado** si puede ser generado por un conjunto finito de vectores. En tal caso se busca un conjunto generador de V que sea "minimal", es decir, que no contenga vectores innecesarios o redundantes. Por los teoremas **4.6.1** y **4.7.1**, esto implica que tal conjunto debe ser **linealmente independiente**. Un conjunto generador de este tipo se denomina **base** del espacio vectorial:

Un conjunto de vectores $B = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\} \subset V$ es una ***base*** del espacio vectorial V si y sólo si:
1) Es linealmente independiente
2) Genera V: $V = \text{gen}(B) = \langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \rangle$.

El número n de elementos de la base es la **dimensión** del espacio vectorial V. Se lo indica como $\dim V = n$.
Demostraremos luego que todas las bases de un espacio vectorial V finitamente generado tienen el **mismo** número de elementos).

Para completar la definición anterior, el subespacio nulo $V = \{\mathbf{0}\}$ se dice que tiene dimensión 0 (no contiene vectores linealmente independientes).

Y los espacios vectoriales que no son finitamente generados (tal como el espacio de funciones continuas $C(\mathbb{R})$) se dice que tienen dimensión **infinita**.

Cuando dos espacios tienen la misma dimensión (finita) se dice que son *isomorfos*. Profundizaremos este concepto en el próximo capítulo.

Ejemplos 4.8.1

1) El conjunto $B = \{\boldsymbol{e}_1, \boldsymbol{e}_2\} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \subset \mathbb{R}^2$ es una base de $\mathbb{R}^2$ (que aquí identificamos con $\mathbb{R}^{2\times 1}$) denominada **base canónica**. Es claro que son linealmente independientes (problema 4.7.2) y que generan $\mathbb{R}^2$ (ejemplo 4.6.1): si $\begin{pmatrix} x \\ y \end{pmatrix}$ es un vector génerico de $\mathbb{R}^2$,

$$\begin{pmatrix} x \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix} = x\boldsymbol{e}_1 + y\boldsymbol{e}_2$$

Por lo tanto $\dim \mathbb{R}^2 = 2$.

2) Analogamente, el conjunto $B = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \subset \mathbb{R}^3$ es la **base canónica** de $\mathbb{R}^3$. Es obvio que son linealmente independientes (¡probar!) y que generan $\mathbb{R}^3$ (ejemplo 4.6.2):

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = x\boldsymbol{e}_1 + y\boldsymbol{e}_2 + z\boldsymbol{e}_3$$

Por lo tanto $\dim \mathbb{R}^3 = 3$

3) Generalizando, el conjunto $B=\{\boldsymbol{e}_1,\boldsymbol{e}_2,\ldots,\boldsymbol{e}_n\}=\left\{\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix},\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}\right\}\subset\mathbb{R}^n$ es la **base canónica** de $\mathbb{R}^n$. Son linealmente independientes (prob. 4.7.2) y generan $\mathbb{R}^n$:

$$\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=x_1\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}+x_2\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}+\ldots+x_n\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}=x_1\boldsymbol{e}_1+\ldots+x_n\boldsymbol{e}_n$$

Por lo tanto $\dim\mathbb{R}^n=n$

4) El conjunto $B=\left\{\begin{pmatrix}1&0\\0&0\end{pmatrix},\begin{pmatrix}0&1\\0&0\end{pmatrix},\begin{pmatrix}0&0\\1&0\end{pmatrix},\ldots,\begin{pmatrix}0&0\\0&1\end{pmatrix}\right\}\subset\mathbb{R}^{2\times 2}$ es la **base canónica** de $\mathbb{R}^{2\times 2}$. En efecto, son linealmente independientes (¡probar!) y generan $\mathbb{R}^{2\times 2}$ (ejemplo 4.6.3):

$$\begin{pmatrix}x_1&x_2\\x_3&x_4\end{pmatrix}=x_1\begin{pmatrix}1&0\\0&0\end{pmatrix}+x_2\begin{pmatrix}0&1\\0&0\end{pmatrix}+x_3\begin{pmatrix}0&0\\1&0\end{pmatrix}+x_4\begin{pmatrix}0&0\\0&1\end{pmatrix}$$

Por lo tanto $\dim\mathbb{R}^{2\times 2}=4$

5) El conjunto de matrices de $m\times n$

$$B=\{E_{11},E_{12},\ldots,E_{mn}\}=\left\{\begin{pmatrix}1&0&\ldots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\ldots&0\end{pmatrix},\begin{pmatrix}0&1&\ldots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\ldots&0\end{pmatrix},\ldots,\begin{pmatrix}0&0&\ldots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\ldots&1\end{pmatrix}\right\}$$

es la base canónica de $V=\mathbb{R}^{m\times n}$. Se deja como ejercicio ver que son linealmente independientes y que generan $\mathbb{R}^{m\times n}$:

$$\begin{pmatrix}x_{11}&x_{12}&\ldots&x_{1n}\\\vdots&\vdots&\ddots&\vdots\\x_{m1}&x_{m2}&\ldots&x_{mn}\end{pmatrix}=x_{11}E_{11}+x_{12}E_{12}+\ldots+x_{mn}E_{mn}$$

Por lo tanto $\dim\mathbb{R}^{m\times n}=m\,n$

6) El conjunto de polinomios $B_n=\{1,t,\ldots,t^n\}$ es la base canónica de P_n. Es linealmente independiente (si $a_0+a_1t+\ldots+a_nt^n=0=0+0t+\ldots+0t^n\ \forall\ t\in\mathbb{R}\Rightarrow a_0=a_1=\ldots=a_n=0$, por igualdad de los coeficientes de las potencias del mismo grado) y todo polinomio $p(t)=a_0+a_1t+\ldots+a_nt^n\in P_n$ es combinación lineal de los elementos de B_n. Por lo tanto $\dim P_n=n+1$.
Observación. El espacio P de **todos** los polinomios tiene dimensión **infinita**, ya que ningún conjunto finito de polinomios puede generarlo.

7) Dado que los subespacios son espacios vectoriales, los conceptos de base y dimensión se aplican también a los mismos. Por ejemplo, sea $A=\begin{pmatrix}1&1&0&1\\2&3&1&2\end{pmatrix}\in\mathbb{R}^{2\times 4}$. Se mostró en el ejemplo 4.4.1 que su espacio nulo es
$N(A)=\left\{x_3\begin{pmatrix}1\\-1\\1\\0\end{pmatrix}+x_4\begin{pmatrix}-1\\0\\0\\1\end{pmatrix},x_3,x_4\in\mathbb{R}\right\}\subset\mathbb{R}^4$. Entonces $B=\left\{\begin{pmatrix}1\\-1\\1\\0\end{pmatrix},\begin{pmatrix}-1\\0\\0\\1\end{pmatrix}\right\}$ es una base de $N(A)$, ya que estos dos vectores 1) son linealmente independientes (no son proporcionales) y 2) generan $N(A)$. Por lo tanto $\dim N(A)=2$.

También existen, por su puesto, otras bases de los conjuntos anteriores. Antes de considerar el caso general, demostremos un teorema para $\mathbb{R}^n$.

Teorema 4.8.1
Todo conjunto linealmente independiente de n vectores de $\mathbb{R}^n$ es una base de $\mathbb{R}^n$.

Demostración.
Sea $M = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ un conjunto linealmente independiente de n vectores de $\mathbb{R}^n$ y sea $\boldsymbol{u}$ un vector aribtrario de $\mathbb{R}^n$. Dado que M es linealmente independiente, para mostrar que es base sólo debemos probar que genera $\mathbb{R}^n$. Consideremos entonces la combinación lineal

$$\alpha_1\boldsymbol{v}_1 + \ldots + \alpha_n\boldsymbol{v}_n = \boldsymbol{u}$$

Utilizando la misma notación del teorema **4.7.2**, esta ecuación conduce al sistema no-homogéneo

$$\alpha_1\begin{pmatrix} v_{11} \\ \vdots \\ v_{n1}\end{pmatrix} + \ldots + \alpha_n\begin{pmatrix} v_{1n} \\ \vdots \\ v_{nn}\end{pmatrix} = \begin{pmatrix} v_{11} & \ldots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{n1} & \ldots & v_{nn}\end{pmatrix}\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n\end{pmatrix} = \begin{pmatrix} u_1 \\ \vdots \\ u_n\end{pmatrix}$$

es decir, $A\boldsymbol{\alpha} = \boldsymbol{u}$ con A la matriz formada por los n vectores y $\boldsymbol{\alpha}$ un vector columna de componentes α_i (incógnitas). Por el teorema **4.7.2**, A es no singular por ser M linealmente independiente, por lo que el sistema anterior es siempre **compatible determinado**, con solución **única** $\boldsymbol{\alpha} = A^{-1}\boldsymbol{u}$. Por lo tanto todo vector $\boldsymbol{u}$ puede ser escrito como combinación lineal de los $\boldsymbol{v}_i$, es decir, que M genera $\mathbb{R}^n$. También es claro que si el conjunto M es linealmente dependiente, la matriz A será singular (teorema **4.7.2**) y en tal caso el sistema anterior no será siempre compatible, es decir, M no podrá generar todo $\mathbb{R}^n$.

Ejemplos 4.8.2

1) El conjunto $\left\{\begin{pmatrix} 1 \\ 2\end{pmatrix}, \begin{pmatrix} 2 \\ 1\end{pmatrix}\right\}$ es una base de $\mathbb{R}^2$ pues son 2 vectores linealmente independientes.

1) El conjunto $\left\{\begin{pmatrix} 1 \\ 0 \\ 0\end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0\end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1\end{pmatrix}\right\}$ es una base de $\mathbb{R}^3$ pues son 3 vectores linealmente independientes ($\det A = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1\end{vmatrix} = 1 \neq 0$).

Es claro también que cualquier conjunto de $m > n$ vectores de $\mathbb{R}^n$ no puede ser base de $\mathbb{R}^n$ pues por el teorema **4.7.3** son linealmente dependientes. Tampoco lo puede ser un conjunto de $m < n$ vectores, porque aun si es linealmente independiente, no generará todo $\mathbb{R}^n$ (¡justificar!, planteando el sistema correspondiente).

Para extender los resultados anteriores a un espacio vectorial V general, demostraremos el siguiente teorema, que generaliza el teorema **4.7.3**

Teorema 4.8.2
Si $B = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ es una base de un espacio vectorial V, todo conjunto $M = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m\}$ de $m > n$ vectores de V es linealmente dependiente.

Demostración.
Dado que B es base de V, todo vector de M es combinación lineal de los $\boldsymbol{v}_i$:

$$\boldsymbol{u}_i = \beta_{1i}\boldsymbol{v}_1 + \ldots + \beta_{ni}\boldsymbol{v}_n, \quad i = 1, \ldots, m$$

Reemplazando esta expresión en la combinación lineal

$$\alpha_1\boldsymbol{u}_1 + \ldots + \alpha_m\boldsymbol{u}_m = \mathbf{0}$$

obtenemos

$$\alpha_1(\beta_{11}\boldsymbol{v}_1 + \ldots + \beta_{n1}\boldsymbol{v}_n) + \ldots + \alpha_m(\beta_{1m}\boldsymbol{v}_1 + \ldots + \beta_{nm}\boldsymbol{v}_n) = \mathbf{0}$$

es decir,

$$(\beta_{11}\alpha_1 + \ldots + \beta_{1m}\alpha_m)\boldsymbol{v}_1 + \ldots + (\beta_{n1}\alpha_1 + \ldots + \beta_{nm}\alpha_m)\boldsymbol{v}_n = \mathbf{0}$$

Pero como los $\boldsymbol{v}_i$ son linealmente independientes, todos los coeficientes deben ser nulos:

$$\begin{array}{rcl} \beta_{11}\alpha_1 + \ldots + \beta_{1m}\alpha_m & = & 0 \\ \vdots & \vdots & \vdots \\ \beta_{n1}\alpha_1 + \ldots + \beta_{nm}\alpha_m & = & 0 \end{array}$$

Este es un sistema de n ecuaciones homogéneas con $m > n$ incógnitas $\alpha_1, \ldots, \alpha_m$, siendo entonces compatible indeterminado, con infinitas soluciones no triviales para los α_i. Por lo tanto, $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m\}$ es linealmente dependiente.

Como consecuencia de este teorema, tenemos el fundamental corolario siguiente:

Corolario 4.8.2
Si $B = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ y $B' = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m\}$ son dos bases de un espacio vectorial V $\Rightarrow m = n$.
Es decir, **todas las bases de un espacio vectorial V finitamente generado tienen el mismo número de elementos**.

Demostración.
Como B y B' son bases, son conjuntos linealmente independientes. Pero si $m > n$, B' sería linealmente dependiente, por el teorema anterior, por lo que necesariamente $m \leq n$. Análogamente, si $m < n$ y B' es base, el conjunto B sería linealmente dependiente por el mismo teorema anterior. Por lo tanto, necesariamente $m = n$. Este resultado permitirá definir la **dimensión** de un espacio vectorial como el número de elementos de cualquier base del mismo.

Como segunda consecuencia, se obtiene la generalización del teorema **4.8.1**:

Teorema 4.8.3
Si un espacio de V tiene dimensión n, todo conjunto $M = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ linealmente independiente de n vectores de V es una base de V.

Demostración.
Dado que el conjunto es linealmente independiente, solo debemos probar que genera V.

Sea $\boldsymbol{u}$ un vector cualquiera de V. Por el teorema anterior, el conjunto de $n+1$ vectores $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n, \boldsymbol{u}\}$ es linealmente dependiente, por lo que existen coeficientes $\alpha_1, \ldots, \alpha_n, \beta$ no todos nulos tales que

$$\alpha_1\boldsymbol{v}_1 + \ldots + \alpha_n\boldsymbol{v}_n + \beta\boldsymbol{u} = \boldsymbol{0}$$

Si $\beta = 0$, la ecuación anterior se reduce a $\alpha_1\boldsymbol{v}_1 + \ldots + \alpha_n\boldsymbol{v}_n = \boldsymbol{0}$, que implica $\alpha_1 = \ldots = \alpha_n = 0$ por la independencia lineal de los $\boldsymbol{v}_i$. Como no son todos nulos, necesariamente $\beta \neq 0$, por lo que se puede despejar $\boldsymbol{u}$ como

$$\boldsymbol{u} = -(\alpha_1\boldsymbol{v}_1 + \ldots + \alpha_n\boldsymbol{v}_n)/\beta = (-\tfrac{\alpha_1}{\beta})\boldsymbol{v}_1 + \ldots + (-\tfrac{\alpha_n}{\beta})\boldsymbol{v}_n$$

Esto muestra que $\boldsymbol{u} \in \text{gen}(M) = \langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \rangle$, es decir, que M genera V. Este teorema generaliza el teorema **4.8.1** a todo espacio de dimensión n.

Otras propiedades importantes para un espacio vectorial de dimensión n que se derivan de los dos teoremas anteriores son:
I. Todo conjunto linealmente independiente de $m < n$ vectores de V puede ser extendido para formar una base de V.
II. Cualquier conjunto con $m > n$ vectores de V que genere V puede ser recortado para formar una base de V.
Se deja la demostración de estos dos enunciados como problema.

Ejemplos 4.8.3

1) Sea $V = \mathbb{R}^3$ y $\boldsymbol{v} \in \mathbb{R}^3$ un vector **no nulo**. Entonces el espacio generado por $\boldsymbol{v}$ es un subespacio de $\mathbb{R}^3$ de dimensión 1, que geométricamente es una recta que pasa por el origen, siendo $B = \{\boldsymbol{v}\}$ una base del mismo:

$$S = \langle \boldsymbol{v} \rangle = \{\alpha\boldsymbol{v}, \alpha \in \mathbb{R}\}, \ \ \dim S = 1 \quad (\boldsymbol{v} \neq \boldsymbol{0})$$

Por su puesto, cualquier vector $\alpha\boldsymbol{v}$, con $\alpha \neq 0$, es también base del mismo (¡probar!). Las mismas consideraciones son válidas para el espacio generado por un vector no nulo de $\mathbb{R}^n$.

2) Sea $V = \mathbb{R}^3$ y $\{\boldsymbol{v}, \boldsymbol{w}\} \in \mathbb{R}^3$ un conjunto de dos vectores linealmente independientes (o sea, no nulos y no proporcionales). Entonces el espacio que generan es un subespacio de $\mathbb{R}^3$ de dimensión 2, que geométricamente es un plano que pasa por el origen, siendo $B = \{\boldsymbol{v}, \boldsymbol{w}\}$ una base del mismo:

$$S = \langle \boldsymbol{v}, \boldsymbol{w} \rangle = \{\alpha\boldsymbol{v} + \beta\boldsymbol{w}, \alpha, \beta \in \mathbb{R}\}, \ \ \dim S = 2 \ (\{\boldsymbol{v}, \boldsymbol{w}\} \text{ linealmente indep.})$$

3) Sea $S = \{(x, y, z) \in \mathbb{R}^3, x + y + z = 0\}$. Geométricamente, S es un plano que pasa por el origen, ortogonal al vector $(1, 1, 1)$. Para determinar una base de S y su dimensión (que será obviamente 2), notamos primero que la solución del "sistema" $x + y + z = 0$ es $x = -y - z$, con y y z libres. Por lo tanto,

$$S = \{(-y-z, y, z), y, z \in \mathbb{R}\} = \{y(-1, 1, 0) + z(-1, 0, 1), y, z \in \mathbb{R}\} = \langle (-1, 1, 0), (-1, 0, 1) \rangle$$

Entonces $B = \{(-1, 1, 0), (-1, 0, 1)\}$ es una base de S, ya que estos vectores son linealmente independientes (no son colineales) y generan S. Por lo tanto $\dim S = 2$.

Problemas 4.8

1) Analizar si los siguientes conjuntos son base de $\mathbb{R}^3$:

i) $\left\{\begin{pmatrix} -3 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ -5 \end{pmatrix}, \begin{pmatrix} 4 \\ 2 \\ 3 \end{pmatrix}\right\}$ ii) $\left\{\begin{pmatrix} 4 \\ -2 \\ 5 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \\ 3 \end{pmatrix}\right\}$ iii) $\left\{\begin{pmatrix} 4 \\ 0 \\ -3 \end{pmatrix}, \begin{pmatrix} 3 \\ 5 \\ 2 \end{pmatrix}\right\}$

2) ¿Puede una base de un espacio vectorial V contener al elemento nulo de V?

3) a) Dado $S = \left\{\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : 3y - 4x = 0\right\}$, hallar una base de S y su dimensión, e interpretar S geométricamente. ¿Pertenecen $\boldsymbol{u} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ y $\boldsymbol{v} = \begin{pmatrix} 6 \\ 8 \end{pmatrix}$ a S?
Si pertenecen, escribirlos como combinación lineal de los vectores de la base.

b) Idem a) para $S = \left\{\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : x - y + 2z = 0\right\}$, con $\boldsymbol{u} = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}$, $\boldsymbol{v} = \begin{pmatrix} 1 \\ -3 \\ -2 \end{pmatrix}$.

c) Idem a) para $S = \left\{\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : x - y + 2z = 0, x + y - z = 0\right\}$,
con $\boldsymbol{u} = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}$, $\boldsymbol{v} = \begin{pmatrix} 1 \\ -3 \\ -2 \end{pmatrix}$.

4) Sea S el subespacio de $\mathbb{R}^3$ generado por el conjunto $\{(1,0,1),(1,2,-1),(1,-2,3)\}$. Hallar una base de S y su dimensión (verificar si son linealmente independientes).

5) Sea S el conjunto de ternas $(x,y,z) \in \mathbb{R}^3$ tales que la suma de los dos primeros números es igual al tercero, y la resta de los dos primeros es igual a la mitad del tercero. ¿Es S un subespacio de $\mathbb{R}^3$? En tal caso hallar una base de S y su dimensión.

6) ¿ Para qué valores de a es $B = \left\{\begin{pmatrix} a \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ a \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2-a \end{pmatrix}\right\}$ una base de R^3?

7) Hallar una base del espacio nulo de las matrices dadas e indicar su dimensión:

i) $A = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 3 \\ 4 & 3 & 7 \end{pmatrix}$, ii) $A = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 3 & 3 & 6 & 9 \\ 4 & 4 & 8 & 12 \end{pmatrix}$

8) a) Sea S el subconjunto de polinomios de grado ≤ 2 que satisfacen $p(1) = p(2)$. Indique si S es un subespacio de P_2, y en tal caso halle una base de S y su dimensión.
b) Determine si el conjunto $\{t, 1+t+t^2, 1+t-t^2\}$ es base de P_2.

9) Determinar la dimensión y una base de los siguientes subespacios:
a) El susbespacio de $\mathbb{R}^{2\times 2}$ de matrices simétricas ($A^T = A$).
b) El susbespacio de $\mathbb{R}^{n\times n}$ de matrices simétricas.
c) El susbespacio de $\mathbb{R}^{2\times 2}$ de matrices antisimétricas ($A^T = -A$)
d) El susbespacio de $\mathbb{R}^{n\times n}$ de matrices antisimétricas.

10) Sea $B = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3\}$ una base de un espacio vectorial V real, y sean α_i, β_i reales.
a) Mostrar que $B' = \{\alpha_1\boldsymbol{b}_1, \alpha_2\boldsymbol{b}_2, \alpha_3\boldsymbol{b}_3\}$ es base de V siempre y cuando $\alpha_1\alpha_2\alpha_3 \neq 0$.
b) Mostrar que $B'' = \{\boldsymbol{b}_1, \boldsymbol{b}_2 + \alpha_1\boldsymbol{b}_1, \boldsymbol{b}_3 + \beta_1\boldsymbol{b}_1 + \beta_2\boldsymbol{b}_2\}$ es base de V $\forall$ α_1, β_1, β_2.
Interprete estos resultados geométricamente.

11) Sea V un espacio vectorial de dimensión n. Muestre que:
a) Ningún conjunto con menos de n vectores puede generar V.
b) Un conjunto de n vectores linealmente dependiente no puede generar V.
c) Todo conjunto linealmente independiente de $m < n$ vectores de V puede ser extendido para formar una base de V .
d) Cualquier conjunto con $m > n$ vectores de V que genere V puede ser recortado para formar una base de V.

Comentario
Las bases canónicas parecen ser las más simples y las más naturales para ser utilizadas. Sin embargo, en algunas aplicaciones resulta conveniente utlizar otras bases, como veremos en capítulos posteriores. Esto nos lleva a la necesidad de poder realizar *cambios de base*, tema que se discutirá en la próxima sección.

4.9. Coordenadas de un vector en una base y cambio de base

Sea V un espacio vectorial finitamente generado y $B = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ una base de V. Todo vector $\boldsymbol{u} \in V$ puede escribirse como combinación lineal **única** de los vectores de la base, ya que estos vectores generan V y son linealmente independientes (teorema de unicidad):

$$\boldsymbol{u} = \alpha_1\boldsymbol{v}_1 + \ldots + \alpha_n\boldsymbol{v}_n \tag{4.9.1}$$

Los escalares $\alpha_1, \ldots, \alpha_n$ se denominan **coordenadas del vector $\boldsymbol{u}$ en la base** B. Se los escribe normalmente como un vector columna $[\boldsymbol{u}]_B \in \mathbb{R}^n$:

$$[\boldsymbol{u}]_B = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

En este contexto, se sobreentiende a B como una **base ordenada** $B = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$, tal que los subíndices de los elementos de la base indican el orden de los mismos. Así, α_1 es la coordenada asociada a $\boldsymbol{v}_1$, α_2 la asociada a $\boldsymbol{v}_2$, etc.

En particular, si $V = \mathbb{R}^n$ y B es la base canónica ordenada $B_c = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$, el vector de coordenadas de $\boldsymbol{u} \in \mathbb{R}^n$ en la base canónica es el mismo vector $\boldsymbol{u}$ (escrito como vector columna). Por ejemplo, en $\mathbb{R}^2$,

$$\boldsymbol{u} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = x\boldsymbol{e}_1 + y\boldsymbol{e}_2$$

con $\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, por lo que

$$[\boldsymbol{u}]_{B_c} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Si consideramos ahora una base distinta, por ejemplo $B = (\boldsymbol{v}_1, \boldsymbol{v}_2) = \left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}\right)$ (la cual es base pues son dos vectores linealmente independientes de $\mathbb{R}^2$), tendremos

$$\boldsymbol{u} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 = \alpha_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \alpha_2 \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Para obtener las nuevas coordenadas α_1, α_2, se debe entonces resolver el sistema

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

es decir,

$$A[\boldsymbol{u}]_B = [\boldsymbol{u}]_{B_c}$$

donde $A = ([\boldsymbol{v}_1]_{B_c}, [\boldsymbol{v}_2]_{B_c}) = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ es la matriz de coordenadas de los vectores de la nueva base B en la base canónica. Esta matriz es **no singular** por ser B linealmente independiente (Teorema 4.7.2), existiendo entonces su inversa. El sistema anterior tendrá así la solución única,

$$[\boldsymbol{u}]_B = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} -1 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} -x_1 + 2x_2 \\ 2x_1 - x_2 \end{pmatrix}$$

es decir, $\alpha_1 = \frac{-x_1+2x_2}{3}$, $\alpha_2 = \frac{2x_1-x_2}{3}$. Por lo tanto,

$$\boldsymbol{u} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{-x_1+2x_2}{3} \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \frac{2x_1-x_2}{3} \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Por ejemplo, si $\boldsymbol{u} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, obtenemos $\alpha_1 = 1$, $\alpha_2 = 0$, ya que $\boldsymbol{u} = \boldsymbol{v}_1 = 1\boldsymbol{v}_1 + 0\boldsymbol{v}_2$.

La interpretación gráfica de las nuevas coordenadas se muestra en la figura. Todo vector $\boldsymbol{u}$ del plano $\mathbb{R}^2$ se puede escribir como combinación lineal de los vectores de la base canónica $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{u} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2$, pero también como combinación lineal de los vectores $\boldsymbol{v}_1, \boldsymbol{v}_2$ de cualquier otra base de $\mathbb{R}^2$, o sea, $\boldsymbol{u} = \alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2$, donde $\boldsymbol{v}_1, \boldsymbol{v}_2$ deben ser no nulos y no paralelos.

En el caso general de $\mathbb{R}^n$, para encontrar las nuevas coordenadas $\alpha_1, \ldots, \alpha_n$ reescribimos la ecuación (4.9.1), en forma matricial, tal como en el teorema 4.7.2:

$$\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = \alpha_1 \begin{pmatrix} v_{11} \\ \vdots \\ v_{n1} \end{pmatrix} + \ldots + \alpha_n \begin{pmatrix} v_{1n} \\ \vdots \\ v_{nn} \end{pmatrix} = \begin{pmatrix} v_{11} & \ldots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{n1} & \ldots & v_{nn} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

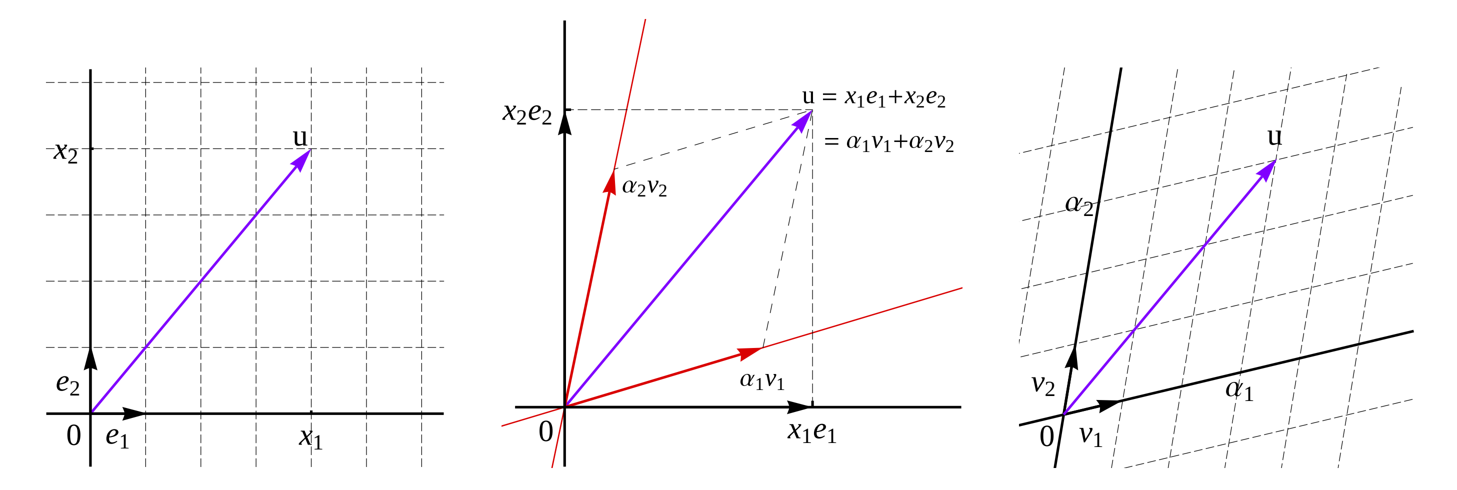


Figura 4.6: Coordenadas de un vector en la base canónica y en una base arbitraria.

o sea,

$$A[\boldsymbol{u}]_B = [\boldsymbol{u}]_{B_c}$$

donde hemos escrito $[\boldsymbol{u}]_{B_c} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ y

$$A = ([\boldsymbol{v}_1]_{B_c}, \ldots, [\boldsymbol{v}_n]_{B_c}) = \begin{pmatrix} v_{11} & \ldots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{n1} & \ldots & v_{nn} \end{pmatrix}$$

es la misma matriz de los teoremas **4.7.2** y **4.8.1**, es decir, la matriz que contiene las coordenadas de los n vectores de la nueva base expresados en la base canónica. Esta matriz es no singular por ser B linealmente independiente (Teorema **4.7.2**). Por lo tanto, las coordenadas en la base B pueden obtenerse como

$$[\boldsymbol{u}]_B = A^{-1}[\boldsymbol{u}]_{B_c} \tag{4.9.2}$$

En este contexto la matriz A se denomina **matriz de cambio de base:**

$$[\boldsymbol{u}]_{B_c} \xrightarrow{A^{-1}} [\boldsymbol{u}]_B$$

La expresión anterior se extiende en forma directa a cualquier espacio vectorial V

finitamente generado:

Teorema 4.9.1 Sea V un espacio vectorial de dimensión finita n y sean $B = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$, $B' = (\boldsymbol{v}'_1, \ldots, \boldsymbol{v}'_n)$ dos bases ordenadas de V. Dado $\boldsymbol{u} \in V$, si

$$\begin{aligned} \boldsymbol{u} &= \alpha_1 \boldsymbol{v}_1 + \ldots + \alpha_n \boldsymbol{v}_n \\ &= \alpha'_1 \boldsymbol{v}'_1 + \ldots + \alpha'_n \boldsymbol{v}'_n \end{aligned}$$

las coordenadas de $\boldsymbol{u}$ en estas bases,

$$[\boldsymbol{u}]_B = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \quad [\boldsymbol{u}]_{B'} = \begin{pmatrix} \alpha'_1 \\ \vdots \\ \alpha'_n \end{pmatrix}$$

satisfacen la ecuación $A[\boldsymbol{u}]_{B'} = [\boldsymbol{u}]_B$, o sea,

$$[\boldsymbol{u}]_{B'} = A^{-1}[\boldsymbol{u}]_B \tag{4.9.3}$$

donde

$$A = ([\boldsymbol{v}'_1]_B, \ldots, [\boldsymbol{v}'_n]_B) = \begin{pmatrix} v'_{11} & \ldots & v'_{1n} \\ \vdots & \ddots & \vdots \\ v'_{n1} & \ldots & v'_{nn} \end{pmatrix}$$

es la matriz de coordenadas de los vectores de la base B' en la base B. Esta matriz es no singular, por ser B y B' conjuntos linealmente independientes.
Así, vemos que A^{-1} es la matriz de cambio de base de B a B', y A la de B' a B.

Demostración. Es similar a la realizada para $\mathbb{R}^n$. Reemplazando $\boldsymbol{v}'_i = v'_{1i}\boldsymbol{v}_1 + \ldots + v'_{ni}\boldsymbol{v}_n$ para $i = 1, \ldots, n$ en la segunda expresión para $\boldsymbol{u}$, obtenemos

$$\begin{aligned} \boldsymbol{u} &= \alpha'_1(v'_{11}\boldsymbol{v}_1 + \ldots + v'_{n1}\boldsymbol{v}_n) + \ldots + \alpha'_n(v'_{1n}\boldsymbol{v}_1 + \ldots + v'_{nn}\boldsymbol{v}_n) \\ &= (v'_{11}\alpha'_1 + v'_{1n}\alpha'_n)\boldsymbol{v}_1 + \ldots + (v'_{n1}\alpha'_1 + \ldots + v'_{nn}\alpha'_n)\boldsymbol{v}_n \end{aligned}$$

Por lo tanto, por unicidad de las coordenadas en la base B, obtenemos el sistema

$$\left\{ \begin{array}{rcl} v'_{11}\alpha'_1 + \ldots + v'_{1n}\alpha'_n &=& \alpha_1 \\ \vdots & \vdots & \vdots \\ v'_{n1}\alpha'_1 + \ldots + v'_{nn}\alpha'_n &=& \alpha_n \end{array} \right.$$

es decir $A[\boldsymbol{u}]_{B'} = [\boldsymbol{u}]_B$, de donde se obtiene la expresión (4.9.3). Nótese también que si $\boldsymbol{u} = \boldsymbol{0}$, entonces $\alpha_1 = \ldots = \alpha_n = 0$, $\alpha'_1 = \ldots = \alpha'_n = 0$, por ser B, B' linealmente independientes, es decir $[\boldsymbol{0}]_B = [\boldsymbol{0}]_{B'} = \boldsymbol{0}$. Esto implica que A debe ser necesariamente no singular, para que la solución trivial $[\boldsymbol{u}]_{B'} = \boldsymbol{0}$ sea la única solución del sistema $A[\boldsymbol{u}]_{B'} = \boldsymbol{0}$.

Ejemplo 4.9.1: **Rotación**.
Un ejemplo corriente de cambio de base es el originado por una rotación del sistema de coordenadas. Dado $\boldsymbol{v} = \binom{x}{y}$, ¿cuales son sus coordenadas en un sistema rotado?

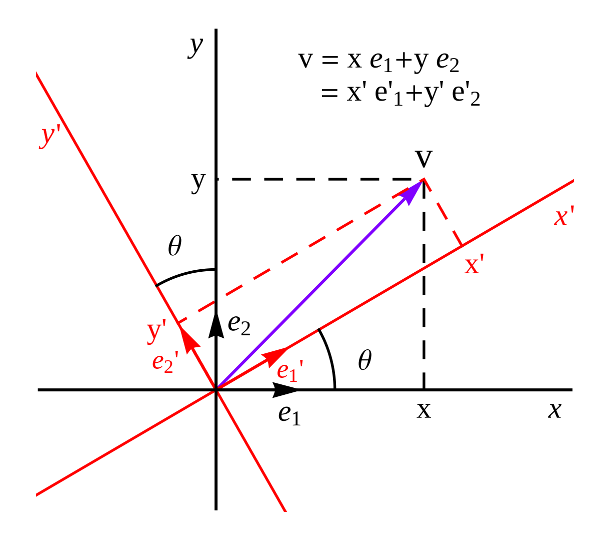


Figura 4.7: Coordenadas de un vector en la base canónica y en la base canónica rotada.

En $V = \mathbb{R}^2$, vemos de la figura que si $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ es la base canónica y $(\boldsymbol{e}'_1, \boldsymbol{e}'_2)$ la base canónica rotada un ángulo θ antihorario, tenemos

$$\boldsymbol{e}'_1 \quad = \quad \begin{pmatrix} \cos\theta \\ \operatorname{sen}\theta \end{pmatrix} = \cos\theta\, \boldsymbol{e}_1 + \operatorname{sen}\theta\, \boldsymbol{e}_2\,, \quad \boldsymbol{e}'_2 = \begin{pmatrix} -\operatorname{sen}\theta \\ \cos\theta \end{pmatrix} = -\operatorname{sen}\theta\, \boldsymbol{e}_1 + \cos\theta\, \boldsymbol{e}_2$$

Por lo tanto, la matriz de cambio de base A toma la forma

$$A = ([\boldsymbol{e}'_1]_{B_c}, [\boldsymbol{e}'_2]_{B_c}) = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

Esta matriz tiene filas y columnas ortonormales, por lo que $A^{-1} = A^T$ (que implica reemplazar θ por $-\theta$, ya que la inversa de una rotación de ángulo θ es una rotación de ángulo $-\theta$). Dado $\boldsymbol{v} = \binom{x}{y} \in \mathbb{R}^2$, podemos escribirlo como

$$\boldsymbol{v} \quad = \quad x\boldsymbol{e}_1 + y\boldsymbol{e}_2 = x'\boldsymbol{e}'_1 + y'\boldsymbol{e}'_2$$

donde las coordenadas en la base rotada estarán dadas por

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = A^{-1} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta + y\operatorname{sen}\theta \\ y\cos\theta - x\operatorname{sen}\theta \end{pmatrix}$$

es decir,

$$x' = x\cos\theta + y\operatorname{sen}\theta, \quad y' = y\cos\theta - x\sin\theta$$

Como aplicación, consideremos el problema de determinar la ecuación de una parábola con vértice en el origen pero rotada un ángulo θ en sentido antihorario (figura 4.8).

Si en el sistema rotado su ecuación es

$$x' = cy'^{\,2}, \quad c > 0$$

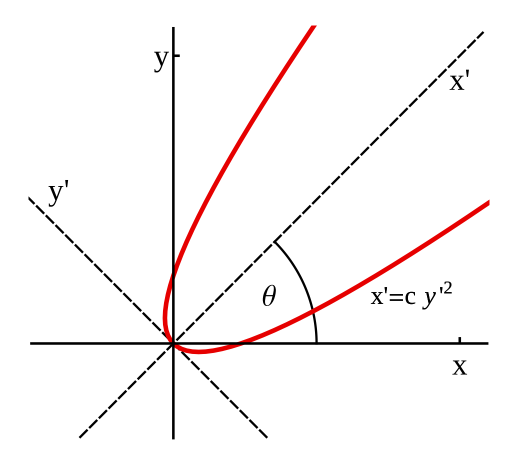


Figura 4.8: Parabola rotada. Mediante cambio de base resulta inmediato escribir la ecuación de la misma.

utlizando las fórmulas anteriores vemos que su ecuación en el sistema original será

$$x\cos\theta + y\,\mathrm{sen}\,\theta = c(y\cos\theta - x\,\mathrm{sen}\,\theta)^2$$

es decir, $x\cos\theta + y\,\mathrm{sen}\,\theta = c(y^2\cos^2\theta + x^2\,\mathrm{sen}^2\,\theta - 2xy\,\mathrm{sen}\,\theta\cos\theta)$.

Ejemplo 4.9.2: Si $\theta = \pi/4$, $\cos\pi/4 = \mathrm{sen}\,\pi/4 = 1/\sqrt{2}$, por lo que

$$\boldsymbol{e}'_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \boldsymbol{e}'_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Por lo tanto, las coordenadas de un vector $\boldsymbol{v} = \binom{x}{y} = x\boldsymbol{e}_1 + y\boldsymbol{e}_2$ en la base rotada son

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} x+y \\ y-x \end{pmatrix}$$

es decir, $x' = \frac{x+y}{\sqrt{2}}$, $y' = \frac{y-x}{\sqrt{2}}$, tal que $\boldsymbol{v} = x'\boldsymbol{e}'_1 + y'\boldsymbol{e}'_2$. La ecuación de la parábola rotada es entonces

$$\sqrt{2}(x+y) = (y-x)^2$$

Problemas 4.9

1) Hallar las coordenadas de $\boldsymbol{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \in \mathbb{R}^2$ en la base $B = \left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$ e interpretar el resultado geométricamente.

2) Hallas las coordenadas de $\boldsymbol{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$ en la base $B = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$.

Hallar en particular las coordenadas de $\boldsymbol{v} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$ e interpretar gráficamente.

3) Encontrar las coordenadas x', y' de un vector $\boldsymbol{v} = \binom{1}{1}$ en un sistema de coordenadas rotado un ángulo de i) 45° antihorario y ii) 30° horario.

4) Determinar la ecuación de una elipse de semieje mayor a y semieje menor b, centrada en el origen y rotada un ángulo θ antihorario, tal que su ecuación en el sistema rotado es $x'^2/a^2 + y'^2/b^2 = 1$. Verificar que si $a = b$ (circunferencia) la ecuación permanece invariante.

5) Considerar una rotación de ángulo θ alrededor del eje y en $\mathbb{R}^3$. Determinar las coordenadas de un vector $\boldsymbol{v} = (x, y, z)^T$ en la base rotada.

4.10. Espacio fila, espacio columna y rango de una matriz

Retornamos ahora al mundo de las matrices y sistemas de ecuaciones. Utilizaremos los conceptos de espacio vectorial, independencia lineal y base para lograr una comprensión más profunda de los sistemas de ecuaciones lineales.

Definiciones. Sea $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \in \mathbb{R}^{m\times n}$.

Vectores columnas de A**:** Son los n vectores $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \in \mathbb{R}^m$, $j = 1, \dots, n$, correspondientes a las columnas de A.

Vectores filas de A**:** Son los m vectores $(a_{i1}, \dots, a_{in}) \in \mathbb{R}^n$, $i = 1, \dots, m$, correspondientes a las filas de A.

Espacio columna de A**:** Es el subespacio de $\mathbb{R}^m$ generado por los n vectores columna de A, o sea el conjunto de todas las combinaciones lineales de las columnas de A:

$$EC(A) = \left\langle \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \dots, \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \right\rangle = \left\{ \alpha_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \dots + \alpha_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}, \alpha_i \in \mathbb{R} \right\}$$

Espacio fila de A**:** Es el subespacio de $\mathbb{R}^n$ generado por las m vectores fila de A, o sea el conjunto de todas las combinaciones lineales de las filas de A:

$$EF(A) = \langle (a_{11}, \dots, a_{1n}), \dots, (a_{m1}, \dots, a_{mn}) \rangle = \{\alpha_1 (a_{11}, \dots, a_{1n}) + \dots + \alpha_m (a_{m1}, \dots, a_{mn}), \alpha_i \in \mathbb{R}\}$$

Ejemplo 4.10.1 Sea $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}$. El espacio fila de A es

$$EF(A) = \langle (1,2,1), (0,1,1) \rangle = \{\alpha(1,2,1) + \beta(0,1,1), \alpha, \beta \in \mathbb{R}\}$$

que es un subespacio de $\mathbb{R}^3$ de dimensión 2, pues las filas son linealmente independientes (forman una base del espacio que generan). Geométricamente, corresponde a un plano que pasa por el origen.

Las columnas de A son en cambio linealmente dependientes. El espacio columna es

$$\begin{aligned} EC(A) &= \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\rangle = \left\{ \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \alpha, \beta \in \mathbb{R} \right\} \\ &= \mathbb{R}^2 \end{aligned}$$

es decir, es todo $\mathbb{R}^2$, ya que las dos primeras columnas son linealmente independientes y por lo tanto ya generan $\mathbb{R}^2$. Su dimensión es también 2.

Rango de una matriz. La igualdad de las dimensiones de los espacios fila y columna en la matriz del ejemplo anterior no es una casualidad, sino una consecuencia del siguiente teorema, que demostraremos luego.

Teorema 4.10.1. Las dimensiones del espacio fila y del espacio columna de una matriz A de $m \times n$ son **iguales**:

$$\dim EC(A) = \dim EF(A) = r(A)$$

La dimensión $r(A)$ del espacio fila o columna de la matriz se denomina **rango** de la matriz.

Nótese que los espacios fila y columna de una matriz no son necesariamente iguales, tal como se vió en el ejemplo anterior. Sólo sus dimensiones son siempre iguales.

El rango de una matriz es pues el "número de filas linealmente independientes" o el "número de columnas linealmente independientes", los cuales siempre coinciden. El rango no puede entonces superar al mínimo entre el número de filas m y el de columnas n:

$$0 \leq r(A) \leq \text{Min}[m, n]$$

Por ejemplo, si $A \in \mathbb{R}^{2\times 4} \Rightarrow r(A) \leq 2$.

Si el rango es igual al número de filas m, entonces $m \leq n$ y las n columnas generan todo $\mathbb{R}^{m\times 1}$, ya que $\dim EC(A) = m$. Si el rango es igual al número de columnas n, entonces $m \geq n$ y las m filas generan todo $\mathbb{R}^{1\times n}$, ya que $\dim EF(A) = n$. Resumiendo,

Si $r(A) = m$ (número de filas) $\Rightarrow m \leq n$ y $EC(A) = \mathbb{R}^{m\times 1}$.
Si $r(A) = n$ (número de columnas) $\Rightarrow m \geq n$ y $EF(A) = \mathbb{R}^{1\times n}$.

Podemos verificar el teorema **4.10.1** en el caso particular de matrices cuadradas no singulares, a partir de los teoremas previos **4.7.2** y **4.7.3**:

Corolario 4.10.1
Sea $A \in \mathbb{R}^{n\times n}$. Si A es no singular ($\det(A) \neq 0$) tanto las n columnas de A como las n filas de A son conjuntos **linealmente independientes**, por lo cual

$$\det A \neq 0 \Rightarrow r(A) = n$$

con $EF(A) = \mathbb{R}^{1\times n}$, $EC(A) = \mathbb{R}^{n\times 1}$.
En cambio, si A es singular ($\det(A) = 0$) tanto las n columnas de A como las n filas de A son conjuntos **linealmente dependientes**, por lo cual

$$\det A = 0 \Rightarrow r(A) < n$$

Demostración. Por el teorema **4.7.2**, si A es no singular las n columnas de A son linealmente independientes, formando entonces una base de $\mathbb{R}^{n\times 1}$ según el teorema **4.8.1**. Por lo tanto $\dim EC(A) = n$, con $EC(A) = \mathbb{R}^{n\times 1}$.
Como $\det(A^T) = \det(A)$, los mismos resultados son válidos para las filas de A, ya que las columnas de la matriz traspuesta A^T son las filas de A. Por lo tanto, las n filas son también linealmente independientes, formando entonces una base de $\mathbb{R}^{1\times n}$, por lo que $\dim EF(A) = n = \dim EC(A)$, con $EF(A) = \mathbb{R}^{n\times 1}$.

Por el contrario, si $\det(A) = 0$ las columnas de A son linealmente dependientes (teorema **4.7.2**), por lo que $\dim EC(A) < n$. Como esto implica $\det(A^T) = 0$, las filas de A son también linealmente dependientes y entonces $\dim EF(A) < n$.

Por ejemplo, si $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 1 \end{pmatrix}$, $\det A = -1 \neq 0$, por lo cual $r(A) = 3$. Tanto las filas como las columnas de A son linealmente independientes, generando entonces $\mathbb{R}^3$ (identficando $\mathbb{R}^{3\times 1}$ y $\mathbb{R}^{1\times 3}$ con $\mathbb{R}^3$).

Matrices escalonadas reducidas.
Resulta también fácil verificar el teorema **4.10.1** en el caso de una matriz **escalonada reducida** U. Por ej., si U es de la forma

$$U = \begin{pmatrix} 1 & u_{12} & u_{13} & u_{14} & u_{15} \\ 0 & 1 & u_{23} & u_{24} & u_{25} \\ 0 & 0 & 0 & 1 & u_{25} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad (4.10.1)$$

se observa que:
1) La filas **no nulas** de U son **linealmente independientes** (ninguna puede escribirse como combinación lineal de las restantes), constituyendo entonces una **base** de $EF(U)$.
2) La columnas con un elemento **"pivote"** (la 1, 2 y 4 en 4.10.1) son **linealmente independientes** y constituyen una base de $EC(U)$, ya que las restantes columnas (3 y 5) pueden escribirse como combinación lineal de estas columnas.
3) El número de filas no nulas y el número de columnas con pivotes **es el mismo** (3 en este caso), ya que cada elemento pivote está asociado a una fila (y columna) distinta.
Entonces **1)+2)+3)** implican

$$\dim EF(U) = \dim EC(U) = r(U)$$

siendo $r(U)$ **el número de filas no nulas** de U, o equivalentemente, el número de columnas de U con un elemento pivote.
Estas consideraciones se aplican a toda matriz escalonada reducida.

Para encontrar una base del espacio fila en el caso general, es útil el siguiente teorema:

Teorema 4.10.2. Sean A y B dos matrices $\in \mathbb{R}^{m\times n}$. Si A y B son equivalentes por filas $\Rightarrow$ tienen el mismo espacio fila:

$$EF(A) = EF(B)$$

Remarquemos, no obstante, que el espacio columna de B **no** es necesariamente igual al espacio columna de A.

Demostración. Si B es equivalente por filas a A, B se obtiene de A por medio de una secuencia finita de operaciones elementales por filas (permutación de filas, multiplicación

de una fila por un escalar no nulo y suma a una fila de otra fila multiplicada por una constante), por lo que las filas de B son combinaciones lineales de las filas de A y por ende $EF(B) \subset EF(A)$. Como A es también equivalente por filas a B (todas las operaciones por fila son invertibles) tenemos $EF(A) \subset EF(B)$. Por lo tanto, $EF(A) = EF(B)$.
Por otro lado, las columnas de B no serán en general combinaciones lineales de las de A. Por ejemplo, $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ y $B = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ son equivalente por filas (están relacionadas por una permutación de filas), pero el espacio columna es claramente distinto.

Corolario 4.10.2 Si U es la matriz escalonada reducida obtenida a partir de A por operaciones elementales de fila, las filas no nulas de U forman una base del espacio fila de A.

Demostración. Por ser U escalonada reducida, las filas no nulas de U son una base del espacio fila de U, que por el teorema anterior, coincide con el espacio fila de la matriz original A. Por lo tanto, son también base de $EF(A)$.

Y para encontrar una base del espacio columna, existen dos formas: Una es encontrar una base del espacio fila de la traspuesta A^T por el procedimiento anterior, ya que $EF(A^T) = EC(A)$ (identificando vectores columna con vectores fila de la misma longitud). La otra, más fácil ya que no requiere una nueva reducción, es utilizar el siguiente resultado, que también demostraremos luego junto con el teorema **4.10.1**:

Corolario 4.10.3 Si U es la matriz escalonada reducida obtenida a partir de A por operaciones elementales de fila, las columnas de la matriz original A correspondientes a las columnas con pivotes de U son una base del espacio columna de la matriz original A.

Remarcamos que mientras las filas no nulas de U son una base del espacio fila de A, las columnas con pivotes de U no son en general una base del espacio columna de A. Sí lo son, en cambio, las correspondientes columnas de A (por ejemplo, las columnas 1, 2 y 4 de A si U tiene la forma (4.10.1)).

Ejemplo 4.10.2

Consideremos $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 1 & 1 \\ 1 & 4 & -1 & -2 \end{pmatrix}$. Reduciéndola por filas, obtenemos:

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 1 & 1 \\ 1 & -4 & -1 & -2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -2 & -3 \\ 0 & 2 & -4 & -6 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix} = U$$

Por lo tanto, $r(A) = 2$ (número de filas no nulas de la última matriz), siendo

$$B_F = \{(1, 2, 3, 4), (0, 1, -2, 3)\}$$

(las filas no nulas de U) una base del espacio fila $EF(A)$ y

$$B_C = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} \right\}$$

una base del espacio columna $EC(A)$ (son las columnas 1 y 2 de A, que corresponden a las columnas con pivote de U). La dimensión de $EF(A)$ y $EC(A)$ es entonces 2.

Comentarios: Las bases de $EF(A)$ y $EC(A)$ no son, por supuesto, únicas. Por ejemplo, puede también llevarse A a la forma escalonada reducida de Gauss-Jordan,

$$A=\begin{pmatrix}1&2&3&4\\1&3&1&1\\1&4&-1&-2\end{pmatrix}\longrightarrow\ldots\longrightarrow\begin{pmatrix}1&2&3&4\\0&1&-2&-3\\0&0&0&0\end{pmatrix}\longrightarrow\begin{pmatrix}1&0&7&10\\0&1&-2&-3\\0&0&0&0\end{pmatrix}=U'$$

y en tal caso $B'_F=\{(1,0,7,10),(0,1,-2,-3)\}$ es también una base de $EF(A)$.

Además, dado que $\dim EF(A)=\dim EC(A)=2$, cualquier par de vectores fila linealmente independientes $\in EF(A)$ forman una base de $EF(A)$ y similarmente, cualquier par de vectores columna linealmente independientes $\in EC(A)$ forman también una base de $EC(A)$. Por ejemplo, $B'_F=\{(1,2,3,4),(1,3,1,1)\}$ es también una base de $EF(A)$, pues $\in EF(A)$ y son linealmente independientes, y
$B'_C=\left\{\begin{pmatrix}1\\1\\1\end{pmatrix},\begin{pmatrix}3\\1\\-1\end{pmatrix}\right\}$, $B''_C=\left\{\begin{pmatrix}2\\3\\4\end{pmatrix},\begin{pmatrix}4\\1\\-2\end{pmatrix}\right\}$ son también bases de $EC(A)$, dado que $\in EC(A)$ y son linealmente independientes.

Ejemplo 4.10.3 Encontrar la dimensión y una base del espacio S generado por el conjunto de vectores $M=\{(1,2,3,0),(1,3,1,1),(1,4,-1,2),(0,1,-2,1)\}$.

El método estándar es entonces formar una matriz A con los vectores puestos por fila y realizar la reducción por filas de la matriz resultante. Obviamente, si fuesen linealmente independientes generarían $\mathbb{R}^4$, pero este no es necesariamente el caso:

$$A=\begin{pmatrix}1&2&3&0\\1&3&1&1\\1&4&-1&2\\0&1&-2&1\end{pmatrix}\longrightarrow\begin{pmatrix}1&2&3&0\\0&1&-2&1\\0&2&-4&2\\0&1&-2&1\end{pmatrix}\longrightarrow\begin{pmatrix}1&2&3&0\\0&1&-2&1\\0&0&0&0\\0&0&0&0\end{pmatrix}=U$$

Esto implica que los 4 vectores generan en realidad un subespacio de $\mathbb{R}^4$ de dimensión 2, siendo $B=\{(1,2,3,0),(0,1,-2,1)\}$ una base del mismo.
Todo vector $\boldsymbol{v}\in S$ puede entonces escribirse como

$$\boldsymbol{v}=\alpha(1,2,3,0)+\beta(0,1,-2,1)$$

con $\alpha,\beta\in\mathbb{R}$.

4.11. Teorema Rango-Nulidad

Recordemos que el **espacio nulo** de una matriz A (sección 4.4) es el conjunto de soluciones del sistema homogéneo asociado: $N(A) = \{\boldsymbol{v} \in \mathbb{R}^n, A\boldsymbol{v} = \mathbf{0}\}$.

Se denomina **nulidad** de una matriz A a la **dimensión** de su espacio nulo $N(A)$:

$$n(A) = \dim N(A)$$

Ejemplo 4.11.1

Consideremos nuevamente la matriz del ejemplo 4.10.2, $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 1 & 1 \\ 1 & 4 & -1 & -2 \end{pmatrix} \in \mathbb{R}^{3\times 4}$.
Para resolver el sistema $A\boldsymbol{v} = \mathbf{0}$ realizamos la reducción por filas de la matriz ampliada,

$$(A|0) = \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 0 \\ 1 & 3 & 1 & 1 & 0 \\ 1 & 4 & -1 & -2 & 0 \end{array}\right) \longrightarrow \ldots \longrightarrow \left(\begin{array}{cccc|c} 1 & 0 & 7 & 10 & 0 \\ 0 & 1 & -2 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) = (U|0)$$

que conduce a la solución $x_1 = -7x_3 - 10x_4$, $x_2 = 2x_3 + 3x_4$, con x_3, x_4 libres:

$$\begin{aligned} N(A) &= \left\{ \begin{pmatrix} -7x_3 - 10x_4 \\ 2x_3 + 3x_4 \\ x_3 \\ x_4 \end{pmatrix}, x_3, x_4 \in \mathbb{R} \right\} = \left\{ x_3 \begin{pmatrix} -7 \\ 2 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -10 \\ 3 \\ 0 \\ 1 \end{pmatrix}, x_3, x_4 \in \mathbb{R} \right\} \\ &= \left\langle \begin{pmatrix} -7 \\ 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -10 \\ 3 \\ 0 \\ 1 \end{pmatrix} \right\rangle \end{aligned}$$

Por lo tanto, $n(A) = 2$, ya que los dos vectores que generan $N(A)$ son linealmente independientes, formando una base de $N(A)$.

Vemos entonces que cada una de las variables libres tiene asociada uno de los vectores que generan $N(A)$, los cuales son, por construcción, **linealmente independientes**.
Por lo tanto,

La nulidad $n(A)$ es el número de variables libres del conjunto solución del sistema homogéneo $A\boldsymbol{x} = \mathbf{0}$.

El resultado anterior es válido para toda matriz A de $m \times n$. Por el teorema **4.4.1**, la nulidad es también el número de variables libres del sistema no homogéne $A\boldsymbol{x} = \boldsymbol{b}$ cuando este es compatible.

Si A es una matriz de $n \times n$ no singular, la única solución del sistema homogéneo $A\boldsymbol{v} = \mathbf{0}$ es la solución trivial $\boldsymbol{v} = \mathbf{0}$, por lo que en este caso $N(A) = \{\mathbf{0}\}$ es el subespacio nulo y la nulidad es 0: $n(A) = 0$.

El rango de la matriz $r(A)$ y la nulidad están relacionados por el siguiente teorema:

Teorema 4.11.1 (Rango-Nulidad).
Para toda matriz A de $m \times n$ se verifica

$$r(A) + n(A) = n$$

es decir, el rango más la nulidad es siempre igual al número de columnas de la matriz.

Demostración. El sistema $A\boldsymbol{v} = \mathbf{0}$ es equivalente al sistema $U\boldsymbol{v} = \mathbf{0}$, donde U es la matriz escalonada reducida por filas derivada de A. Si $r(A) = r \Rightarrow U$ tiene r filas no nulas, existiendo entonces r variables dependientes y $n - r$ variables independientes o libres. Pero la dimensión de $N(A)$ es justamente el número de variables independientes, por lo que $n(A) = n - r$. Por lo tanto,

$$r(A) + n(A) = r + (n - r) = n$$

Así, en el ejemplo anterior 4.11.1, tenemos $r(A) = 2$ y $n(A) = 2$, verificándose que $r(A) + n(A) = 4$.

Ejemplo 4.11.2. Si el sistema reducido final $U\boldsymbol{v} = \mathbf{0}$ es de la forma

$$(U|0) = \left(\begin{array}{ccccc|c} 1 & 0 & u_{13} & 0 & u_{15} & 0 \\ 0 & 1 & u_{23} & 0 & u_{25} & 0 \\ 0 & 0 & 0 & 1 & u_{35} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

entonces posee tres variables dependientes (x_1, x_2, x_4) asociadas a las columnas con pivote, cuyo número es igual al número de filas no nulas de U, es decir al rango $r(U) = 3$, y dos variables independientes (x_3 y x_5), lo que implica $n(U) = 2$. Por lo tanto, se verifica

$$r(U) + n(U) = 3 + 2 = 5$$

El conjunto solución es $x_1 = -u_{13}x_3 - u_{15}x_5$, $x_2 = -u_{23}x_4 - u_{25}x_5$, $x_4 = -u_{45}x_5$, con x_3, x_5 libres, o sea,

$$N(U) = \left\{ x_3 \begin{pmatrix} -u_{13} \\ -u_{23} \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} -u_{15} \\ -u_{25} \\ 0 \\ -u_{35} \\ 1 \end{pmatrix}, x_3, x_5 \in \mathbb{R} \right\} = \left\langle \begin{pmatrix} -u_{13} \\ -u_{23} \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -u_{15} \\ -u_{25} \\ 0 \\ -u_{35} \\ 1 \end{pmatrix} \right\rangle$$

4.11.1. Interpretación geométrica

El teorema rango-nulidad tiene una clara interpretación geométrica. Sea $A \in \mathbb{R}^{m \times n}$. Dado que todo vector $\boldsymbol{v} \in \mathbb{R}^n$ solución de $A\boldsymbol{v} = \mathbf{0}$ es **ortogonal** a todas las filas de A, el espacio nulo $N(A)$ es el conjunto de todos los vectores ortogonales a las filas de A, es

decir, es el **subespacio ortogonal al espacio fila** $EF(A)$ (o sea, el subespacio $EF_{\perp}(A)$ "complementario" al espacio fila). Por lo tanto, la suma de la dimensión del espacio fila, que es el rango $r(A)$, más la dimensión del espacio nulo, que es la nulidad $n(A)$, debe ser la dimensión del espacio completo, n:

$$\dim EF(A) + \dim\, N(A) = n$$

Por ejemplo, si $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow EF(A) = \langle(1,0,0),(0,1,0)\rangle$ corresponde geométricamente al "plano xy", mientras que $N(A) = \{(0,0,z), z \in \mathbb{R}\} = \langle(0,0,1)\rangle$ corresponde geométricamente al "eje z", es decir, al subespacio de $\mathbb{R}^3$ ortogonal al plano xy. Se verifica entonces $r(A) + n(A) = 2 + 1 = 3$.

En el ejemplo **4.11.1**, se verifica que los vectores $(-7,2,1,0)$ y $(-10,3,0,1) \in \mathbb{R}^4$ que generan $N(A)$ son ortogonales a todas las filas de A y por lo tanto de U. Ellos generan entonces el subespacio de $\mathbb{R}^4$ ortogonal al espacio fila $EF(A)$. Dado que $\dim EF(A) = 2$ $\Rightarrow \dim N(A) = 2$, para que su suma sea 4.

Y en el ejemplo **4.11.2** se verifica también que $(-u_{13}, -u_{23}, 1, 0, 0)$ y $(-u_{15}, -u_{25}, 0, -u_{35}, 1)$ $\in \mathbb{R}^5$ son ortogonales a todas las filas de U. Dado que $\dim EF(U) = 3 \Rightarrow \dim N(U) = 2$, para que su suma sea 5.

Finalmente, notemos que si A de $n \times n$ es no singular, los n vectores fila son linealmente independientes, por lo que no existe ningún vector no nulo ortogonal a todas las filas. Así, el único vector ortogonal a todas ellas es el vector nulo $\mathbf{0}$, por lo que $N(A) = \{\mathbf{0}\}$. El subespacio ortogonal a todas las filas tiene entonces dimensión 0 ($n(A) = 0$).

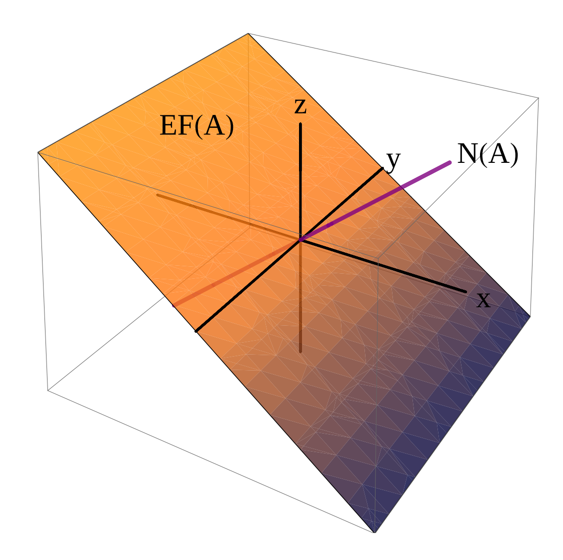


Figura 4.9: Esquema geométrico de un posible espacio fila y espacio nulo de una matriz de 3×3. En la figura, $EF(A)$ es un plano que pasa por el origen ($\dim EF(A) = 2$) y entonces $N(A)$ es la recta perpendicular a $EF(A)$ que pasa por el origen ($\dim N(A) = 1$).

4.12. Aplicación a sistemas de ecuaciones lineales

Aplicaremos ahora los conceptos y resultados anteriores a sistemas de ecuaciones lineales. Sea A una matriz de $m \times n$. Consideremos el sistema asociado de m ecuaciones con n incógnitas

$$A\boldsymbol{x} = \boldsymbol{b}$$

es decir

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

El lado izquierdo puede escribirse como una combinación lineal de columnas de A,

$$x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \ldots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \tag{4.12.1}$$

donde los coeficientes de la combinación lineal son las n incógnitas $x_1, \ldots, x_n$.

A partir de esta expresión se deducen los siguientes teoremas:

Teorema 4.12.1
Sea $A \in \mathbb{R}^{m\times n}$ y $\boldsymbol{b} \in \mathbb{R}^m$. El sistema $A\boldsymbol{x} = \boldsymbol{b}$ es **compatible** si y sólo si $\boldsymbol{b} \in EC(A)$.
(O sea, tiene solución si y sólo si $\boldsymbol{b}$ pertenece al espacio columna de A).

Demostración. Es evidente de la ecuación (4.12.1): Si $\boldsymbol{x}$ es solución del sistema, entonces existe una combinación lineal de las columnas de A que es igual al vector $\boldsymbol{b}$, por lo que $\boldsymbol{b} \in EC(A)$. Y si $\boldsymbol{b} \in EC(A)$, existirá una combinación lineal de las columnas de A que será igual a $\boldsymbol{b}$, por lo que el sistema tendrá solución.

Los sistemas $A\boldsymbol{x} = \boldsymbol{b}$ incompatibles (sin solución) son entonces aquellos en los que $\boldsymbol{b}$ **no pertenece** al espacio columna de la matriz A. Esto puede ocurrir sólo si $EC(A)$ no es todo $\mathbb{R}^m$, sino un subespacio propio de $\mathbb{R}^m$ con dimensión menor que m, o sea, si el rango satisface $r(A) < m$.

Como consecuencia, en los sistemas compatibles **el rango de la matriz ampliada** (formada por A y el vector $\boldsymbol{b}$) **es igual al de la matriz** A (pues $\boldsymbol{b} \in EC(A)$ y entonces el espacio columna de la ampliada coincide con el de A) mientras que en los incompatibles los rangos difieren (pues $\boldsymbol{b} \notin EC(A)$).

Ejemplo 4.12.1: Consideremos el sistema $A\boldsymbol{x} = \boldsymbol{b}$, con $A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 1 & 0 \end{pmatrix}$ y $\boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$. Reduciendo por filas obtenemos

$$\left(\begin{array}{cc|c} 1 & 1 & b_1 \\ 2 & 1 & b_2 \\ 1 & 0 & b_3 \end{array}\right) \longrightarrow \left(\begin{array}{cc|c} 1 & 1 & b_1 \\ 0 & -1 & b_2 - 2b_1 \\ 0 & -1 & b_3 - b_1 \end{array}\right) \longrightarrow \left(\begin{array}{cc|c} 1 & 0 & b_2 - b_1 \\ 0 & 1 & 2b_1 - b_2 \\ 0 & 0 & b_3 - b_2 + b_1 \end{array}\right)$$

En este caso $r(A) = 2 < m = 3$, y existe solución sólo si $b_3 - b_2 + b_1 = 0$, o sea sólo si $b_3 = b_2 - b_1$, que es justamente la condición para que $\boldsymbol{b} \in EC(A) = \left\langle \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle$. El mismo puede ser escrito también como $EC(A) = \left\{ \begin{pmatrix} x \\ y \\ y - x \end{pmatrix}, x, y \in \mathbb{R} \right\}$ (¡probar!).
Si $b \in EC(A)$, la solución única del sistema es $x_1 = b_2 - b_1$, $x_2 = 2b_1 - b_2$, es decir $\boldsymbol{x} = \begin{pmatrix} b_2 - b_1 \\ 2b_1 - b_2 \end{pmatrix}$, cumpliéndose que $x_1 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ (¡verificar!)

Teorema 4.12.2.
Sea $A \in \mathbb{R}^{m \times n}$. El sistema $A\boldsymbol{x} = \boldsymbol{b}$ **es compatible** $\forall\, \boldsymbol{b} \in \mathbb{R}^m$ si y sólo si las columnas de A generan $\mathbb{R}^m$ ($EC(A) = \mathbb{R}^m$), es decir, si y sólo si el rango de A es m ($r(A) = m$).

Demostración. Este caso puede darse sólo si $m \leq n$, es decir, si el número de filas es menor o igual que el número de columnas, para que el rango pueda ser igual al número de filas m. Si $r(A) = m$ entonces $\dim EC(A) = m$, por lo que las n columnas **generan** $\mathbb{R}^m$ ($EC(A) = \mathbb{R}^m$). Por lo tanto, todo vector $\boldsymbol{b} \in \mathbb{R}^m$ podrá ser escrito como combinación lineal de las columnas de A, por lo que existirá solución de $A\boldsymbol{x} = \boldsymbol{b}\ \forall\, \boldsymbol{b} \in \mathbb{R}^m$. Y si $A\boldsymbol{x} = \boldsymbol{b}$ tiene solución $\forall\, \boldsymbol{b} \in \mathbb{R}^m$, entonces todo vector $\boldsymbol{b} \in \mathbb{R}^m$ puede ser escrito como combinación lineal de las columnas de A, por lo que estas generan $\mathbb{R}^m$ ($EC(A) = \mathbb{R}^m$), en cuyo caso $r(A) = \dim EC(A) = m$.

Ejemplo 4.12.2: Consideremos el sistema $A\boldsymbol{x} = \boldsymbol{b}$, con $A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ y $\boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$. Reduciendo por filas obtenemos

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & b_1 \\ 1 & 1 & 0 & b_2 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 2 & 1 & b_1 \\ 0 & 1 & 1 & b_1 - b_2 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 0 & -1 & 2b_2 - b_1 \\ 0 & 1 & 1 & b_1 - b_2 \end{array}\right)$$

En este caso $r(A) = 2 = m$, por lo que $EC(A) = \mathbb{R}^2$, existiendo entonces solución $\forall \boldsymbol{b} \in \mathbb{R}^2$. El conjunto solución es $x_1 = 2b_2 - b_1 + x_3$, $\ x_2 = b_1 - b_2 - x_3$, con x_3 libre:

$$\boldsymbol{x} = \left\{ \begin{pmatrix} 2b_2 - b_1 \\ b_1 - b_2 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, x_3 \in \mathbb{R} \right\},$$

verificándose que $x_1\binom{1}{1} + x_2\binom{2}{1} + x_3\binom{1}{0} = \binom{b_1}{b_2}$ (¡probar!). Nótese que la nulidad es $n(A) = 3 - r(A) = 1$. El espacio nulo es justamente el segundo término en la solución anterior:

$$N(A) = \left\{ x_3 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, x_3 \in \mathbb{R} \right\} = \left\langle \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\rangle$$

Teorema 4.12.3.
Sea $A \in \mathbb{R}^{m \times n}$. El sistema $A\boldsymbol{x} = \boldsymbol{b}$ tiene a lo sumo una solución (o sea, tiene solución única o no tiene ninguna) para todo $\boldsymbol{b} \in \mathbb{R}^m$ si y sólo si las n columnas de A son linealmente independientes, es decir, si y sólo si el rango de A es n ($r(A) = n$)

Demostración. Este caso sólo puede darse si $m \geq n$, para que el rango pueda ser n. Si el sistema $A\boldsymbol{x} = \boldsymbol{b}$ tiene a lo sumo una solución, entonces el sistema homogéneo $A\boldsymbol{x} = \mathbf{0}$ tendrá como única solución la solución trivial $\boldsymbol{x} = \mathbf{0}$, lo que implica, según (4.12.1), que las n columnas son linealmente independientes, es decir, $r(A) = \dim EC(A) = n$.

Y si $r(A) = n \Rightarrow \dim EC(A) = n$, por lo que las n columnas son linealmente independientes. En este caso todo vector $\boldsymbol{b} \in EC(A)$ se escribe de forma **única** como combinación lineal de estas columnas (Teorema **4.7.4**), por lo que el sistema $A\boldsymbol{x} = \boldsymbol{b}$ tendrá solución única $\forall\ \boldsymbol{b} \in EC(A)$. Y no tendrá solución si $\boldsymbol{b} \notin EC(A)$.

Ejemplo 4.12.3
Consideremos primero el sistema del ejemplo **4.12.1**. Las dos columnas de A son linealmente independientes, con $r(A) = n = 2 < m$. Se verifica que el sistema o bien tiene solución única (cuando $\boldsymbol{b} \in EC(A)$, o sea cuando $b_3 = b_2 - b_1$), o bien es incompatible (cuando $\boldsymbol{b} \notin EC(A)$, o sea, $b_3 \neq b_2 - b_1$). Nótese que la nulidad es $n(A) = n - r(A) = 0$.

Consideremos ahora el sistema del ejemplo **4.12.2**. Aquí las columnas de A son **linealmente dependientes**, con $r(A) = 2 < n = 3$, y se verifica que el sistema es compatible **indeterminado**, es decir, no tiene solución única. Nótese que la nulidad es $n(A) = n - r(A) = 1$, indicando que el conjunto solución tendrá un parámetro libre.

Demostración del teorema 4.10.1

Estamos ahora en condiciones de demostrar la igualdad de las dimensiones de los espacios fila y columna. Sea A una matriz de $m \times n$. Si $\dim EF(A) = r$, la matriz U escalonada reducida por filas de A tiene r filas no nulas y entonces r coeficientes 1 como pivotes, siendo las r columnas que contienen estos pivotes **linealmente independientes**. Estas columnas forman una base del espacio columna de U, pero en general no del espacio columna de A, ya que los espacios columna de U y A no son necesariamente coincidentes.

Sea U_r la matriz de $m \times r$ obtenida de U borrando las restantes $n - r$ columnas (las asociadas con las variables libres) y A_r la matriz de $m \times r$ que se obtiene de A borrando las mismas columnas. A_r y U_r siguen siendo equivalentes por filas, por lo que sus espacios fila son coincidentes, teniendo entonces la misma dimensión r.

En el ejemplo 4.11.1, en el que $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 1 & 1 \\ 1 & 4 & -1 & -2 \end{pmatrix}$, $U = \begin{pmatrix} 1 & 0 & 7 & 10 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$,
obtenemos

$$A_r = \begin{pmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix}, \quad U_r = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Los sistemas $A_r \boldsymbol{x} = \mathbf{0}$ y $U_r \boldsymbol{x} = \mathbf{0}$ tienen el mismo conjunto solución, por ser A_r y U_r equivalentes por filas, que es la solución trivial $\boldsymbol{x} = \mathbf{0}$ ya que las r columnas de U_r son linealmente independientes (Teorema 4.12.3). Esto implica, nuevamente por el teorema **4.12.3**, que las r columnas de A_r son linealmente independientes. Por lo tanto, $\dim EC(A_r) = r$ y entonces $\dim EC(A) \geq \dim EC(A_r) = r = \dim EF(A)$, o sea,

$$\dim EC(A) \geq \dim EF(A)$$

Aplicando el mismo razonamiento a la matriz traspuesta A^T, se obtendría
$\dim EC(A^T) \geq \dim EF(A^T)$. Pero como $EC(A^T) = EF(A)$ y $EF(A^T) = EC(A)$, esto implica

$$\dim EF(A) \geq \dim EC(A)$$

Por lo tanto, la única posibilidad es que

$$\dim EF(A) = \dim EC(A)$$

Vemos entonces que las r columnas de A_r, que eran linealmente independientes, forman también una base de $EC(A)$, ya que $\dim EC(A) = r$.

4.12.1. Sistemas $n \times n$

En el caso de matrices cuadradas $A \in \mathbb{R}^{n\times n}$, los resultados anteriores implican:

1. Si A es no singular $\Rightarrow \det(A) \neq 0$ y tanto las n filas como las n columnas de A son **linealmente independientes**, por lo que el rango de A es n y la nulidad 0. Ambos conjuntos son entonces base de $\mathbb{R}^n$, por lo que $EF(A) = EC(A) = \mathbb{R}^n$.
En este caso se cumplen simultáneamente los teoremas **4.12.2** y **4.12.3** ($r(A) = m = n$), por lo que el sistema

$$A\boldsymbol{x} = \boldsymbol{b}$$

tiene solución única $\forall\ \boldsymbol{b} \in \mathbb{R}^n$.

2. Si A es singular $\Rightarrow \det(A) = 0$ y tanto las n filas como las n columnas son **linealmente dependientes**, por lo que el rango de A es menor que n. En este caso ni las filas ni las columnas son base de $\mathbb{R}^n$, por lo que $EF(A)$ y $EC(A)$ son subespacios de $\mathbb{R}^n$ de dimensión $< n$, no necesariamente coincidentes. La nulidad es ≥ 1.
El sistema $A\boldsymbol{x} = \boldsymbol{b}$ será por lo tanto incompatible si $\boldsymbol{b} \notin EC(A)$, y compatible indeterminado (infinitas soluciones) si $\boldsymbol{b} \in EC(A)$.

Estos resultados se pueden resumir en la siguiente tabla:

Matrices $A \in \mathbb{R}^{n\times n}$	
A no singular ($\exists A^{-1}$)	A singular ($\nexists A^{-1}$)
$\det A \neq 0$	$\det A = 0$
rango $r(A) = n$	rango $r(A) \leq n-1$
nulidad $n(A) = 0$	nulidad $n(A) \geq 1$
$r(A) + n(A) = n$	
Tanto las n columnas como las n filas de A son linealmente independientes y generan $\mathbb{R}^n$	Tanto las n columnas como las n filas de A son linealmente dependientes y no generan $\mathbb{R}^n$
$EC(A) = EF(A) = \mathbb{R}^n$	$EC(A) \neq \mathbb{R}^n,\ EF(A) \neq \mathbb{R}^n$
Sistema $A\boldsymbol{x} = \boldsymbol{b}$ tiene solución única $\forall\ \boldsymbol{b} \in \mathbb{R}^n$	Sistema $A\boldsymbol{x} = \boldsymbol{b}$ incompatible si $\boldsymbol{b} \notin EC(A)$ y compatible indeterminado si $\boldsymbol{b} \in EC(A)$

4.12.2. Sistemas $m \times n$

En el caso de matrices no cuadradas $A \in \mathbb{R}^{m\times n}$, los teoremas anteriores implican:

1. Si $m < n$ entonces las n columnas son **linealmente dependientes**, ya que el rango de la matriz satisface $r(A) \leq m < n$. Las m filas son linealmente independientes sólo si $r(A) = m$. La nulidad es entonces **no nula**: $n(A) = n - r(A) \geq 1$. Esto implica que el sistema

$$A\boldsymbol{x} = \boldsymbol{b}$$

que posee más incógnitas que ecuaciones (subdeterminado) **no puede tener solución única**. Si $r(A) = m \Rightarrow EC(A) = \mathbb{R}^m$ y el sistema será compatible indeterminado $\forall\ \boldsymbol{b} \in \mathbb{R}^m$. Pero si $r(A) < m$, las columnas no generan $\mathbb{R}^m$, por lo que el sistema será compatible indeterminado cuando $\boldsymbol{b} \in EC(A)$ e incompatible en caso contrario.

2. Si $m > n$ entonces las m filas son **linealmente dependientes**, ya que el rango de la matriz satisface $r(A) \leq n < m$. Las n columnas son linealmente independientes sólo si $r(A) = n$, y **no pueden generar** $\mathbb{R}^m$, ya que $n < m$. Esto implica que el sistema

$$A\boldsymbol{x} = \boldsymbol{b}$$

que posee más ecuaciones que incógnitas (sobredeterminado) **no puede ser compatible** $\forall\ \boldsymbol{b} \in \mathbb{R}^m$. El sistema será compatible sólo cuando $\boldsymbol{b} \in EC(A)$, siendo en tal caso compatible determinado si $r(A) = n$ (columnas linealmente independientes, nulidad 0) y compatible indeterminado si $r(A) < n$ (columnas linealmente dependientes, nulidad ≥ 1).

Estos resultados se pueden resumir en la siguiente tabla:

<table>
<tr><td colspan="2">Matrices $A \in \mathbb{R}^{m\times n}$</td></tr>
<tr><td>$m < n\,,\quad A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$</td><td>$m > n\,,\quad A = \begin{pmatrix} a_{11}\dots a_{1n} \\ \vdots \\ \vdots \\ a_{m1}\dots a_{mn} \end{pmatrix}$</td></tr>
<tr><td>rango $r(A) \leq m$</td><td>rango $r(A) \leq n$</td></tr>
<tr><td>nulidad $n(A) \geq n - m \geq 1$</td><td>nulidad $n(A) \geq 0$</td></tr>
<tr><td colspan="2">$r(A) + n(A) = n$</td></tr>
<tr><td>Las n columnas son linealmente dependientes
Generan $\mathbb{R}^m$ sólo si $r(A) = m$</td><td>Las n columnas no pueden generar $\mathbb{R}^m$
Son linealmente independientes sólo si $r(A) = n$</td></tr>
<tr><td>$EC(A) = \mathbb{R}^m$ sólo si $r(A) = m$, $EF(A) \neq \mathbb{R}^n$</td><td>$EC(A) \neq \mathbb{R}^m$, $EF(A) = \mathbb{R}^n$ sólo si $r(A) = n$</td></tr>
<tr><td>Sistema $A\boldsymbol{x} = \boldsymbol{b}$ no puede tener solución única
Compatible $\forall\ \boldsymbol{b} \in \mathbb{R}^m$ sólo si $r(A) = m$</td><td>Sistema $A\boldsymbol{x} = \boldsymbol{b}$ no puede ser compatible $\forall\ \boldsymbol{b} \in \mathbb{R}^m$
Con solución única sólo si $r(A) = n$ y $\boldsymbol{b} \in EC(A)$</td></tr>
</table>

Problemas 4.12

1) Determinar el espacio fila, el espacio columna, el espacio nulo, el rango y la nulidad de las siguientes matrices. Determinar también una base de los espacios indicados.

a) $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, b) $A = \begin{pmatrix} 1 & -2 \\ 2 & -4 \end{pmatrix}$, c) $A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

d) $A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & -2 \\ 0 & 1 & 3 \end{pmatrix}$, e) $A = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 3 & 4 & 0 & 2 \end{pmatrix}$, f) $A = \begin{pmatrix} 1 & 0 \\ 3 & 0 \\ 4 & 1 \end{pmatrix}$

g) $A = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix} \in \mathbb{R}^{m\times n}$, h) $A = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix} \in \mathbb{R}^{m\times n}$

2) Indicar para qué valores de α pertencerá
i) $(1, \alpha)$ al espacio fila de las matrices a), b) y f) anteriores.
ii) $\binom{1}{\alpha}$ al espacio columna de las matrices a), b) y e) anteriores.
iii) $(1, 2, \alpha)^T$ al espacio columna de las matrices c), d) y f) anteriores

3) A partir de **2)** iii), indique para qué valores de α el sistema $A\boldsymbol{x} = \boldsymbol{b}$, con $\boldsymbol{b} = (1, 2, \alpha)^T$, será compatible determinado, compatible indeterminado e incompatible, si A es la matriz de i) 1 c), ii) 1 d), iii) 1 f).

4) Indique, en caso de que existan, los valores de k y α para los que los siguientes sistemas serán i) compatibles determinados, ii) compatibles indeterminados y iii) incompatibles, hallando la solución en los casos compatibles e indicando en cada caso el rango y la nulidad de la matriz de coeficientes.

a) $\begin{cases} x + y + z &= 1 \\ 2x - y + z &= 2 \\ 5x - 4y + kz &= \alpha \end{cases}$ b) $\begin{cases} x + y + z + t &= 1 \\ 2x - y + z + 3t &= 2 \\ 5x - 4y + kz + 8t &= \alpha \end{cases}$ c) $\begin{cases} x + y + z &= 1 \\ 2x - y + z &= 2 \\ 5x - 4y + kz &= \alpha \\ 3x + 2z &= 3 \end{cases}$

5) Determine una base y la dimensión del espacio generado por los conjuntos: a) $M = \{(1, 2, 0), (2, 3, 1), (1, 3, -1)\}$
b) $M = \{\binom{1}{2}, \binom{-3}{-6}, \binom{0}{0}\}$
c) $M = \{\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\}$
d) $M = \{1 + t, t + t^2, -1 + t + 2t^2\}$

6) Mostrar que el conjunto de vectores columna $(b_1, b_2, b_3)^T$ para los que el sistema

$$\begin{cases} 3x - 2y - 4z &= b_1 \\ x + z &= b_2 \\ 2x - 2y - 5z &= b_3 \end{cases}$$

es compatible es un subespacio de $\mathbb{R}^3$. Encontrar también su dimensión y una base del mismo.

7) a) Probar que dos matrices poseen el mismo espacio fila si y sólo si, luego de llevarla a la forma reducida de Gauss-Jordan, poseen filas no nulas idénticas.
b) Determinar si los espacios generados por los conjuntos $M = \{(1,2,3),(2,1,1),(3,0,-1)\}$ y $M' = \{(3,3,4),(0,3,5),(1,-1,-2)\}$ son iguales.

8) Indique la forma de las matrices A de $m \times n$ que tienen:
i) rango $r(A) = 0$, ii) rango $r(A) = 1$.

9) a) Mostrar que $\forall\ A \in \mathbb{R}^{m\times n}$, el rango de A y A^T son iguales.
b) Probar que si $A \in \mathbb{R}^{m\times n}$, el sistema $A\boldsymbol{x} = \boldsymbol{b}$ es compatible $\forall\ \boldsymbol{b} \in \mathbb{R}^m$ si y sólo si $r(A) = m$.
c) Probar que si $A \in \mathbb{R}^{m\times n}$ y $r(A) = n \Rightarrow$ si el sistema $A\boldsymbol{x} = \boldsymbol{b}$ es compatible, su solución es única.

10) a) Muestre que $\forall$ matriz $A \in \mathbb{R}^{m\times n}$,
i) el rango satisface $r(A) = r(\alpha A)\ \forall\ \alpha \neq 0$.
ii) $r(MA) = r(A)\ \forall$ matriz $M \in \mathbb{R}^{m\times m}$ no singular.
iii) $r(AS) = r(A)\ \forall$ matriz $S \in \mathbb{R}^{n\times n}$ no singular.
b) Muestre que si U escalonada reducida tiene rango r, entonces contiene una submatriz de $r \times r$ con determinante no nulo, mientras que toda submatriz de $m \times m$ con $m > r$ tiene determinante nulo.
c) Muestre que si A, B son dos matrices $\in \mathbb{R}^{m\times n}$, el rango verifica

$$0 \leq r(A+B) \leq r(A) + r(B)$$

11) a) Muestre que si $\boldsymbol{x}_1$ es una solución del sistema de $m \times n$ $A\boldsymbol{x} = \boldsymbol{b} \Rightarrow \alpha\boldsymbol{x}_1$ es una solución del sistema

$$A\boldsymbol{x} = \alpha\boldsymbol{b}$$

b) Muestre que si $\boldsymbol{x}_1$ es una solución del sistema de $m\times n$ $A\boldsymbol{x} = \boldsymbol{b}_1$ y $\boldsymbol{x}_2$ una solución del sistema $A\boldsymbol{x} = \boldsymbol{b}_2 \Rightarrow \alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2$ es una solución del sistema

$$A\boldsymbol{x} = \alpha_1\boldsymbol{b}_1 + \alpha_2\boldsymbol{b}_2$$

c) Discuta la utilidad y significado de las expresiones anteriores. ¿Cuál debe ser el rango y la nulidad de A para que dichas soluciones sean las únicas?

12) Dada $A = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 1 & 2 \\ 1 & 5 & 3 \end{pmatrix}$, $\boldsymbol{b}_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, $\boldsymbol{b}_2 = \begin{pmatrix} 3 \\ 3 \\ 8 \end{pmatrix}$,
a) Determine la solución de los sistemas $A\boldsymbol{x} = \boldsymbol{b}_1$ y $A\boldsymbol{x} = \boldsymbol{b}_2$.
b) A partir de las soluciones anteriores, encuentre la solución de los sistemas
i) $A\boldsymbol{x} = 2\boldsymbol{b}_1$, ii) $A\boldsymbol{x} = 3\boldsymbol{b}_1 - 4\boldsymbol{b}_2$, iii) $3A\boldsymbol{x} = \boldsymbol{b}_1 - \boldsymbol{b}_2$

Capítulo 5

Transformaciones Lineales

5.1. Introducción

Estudiaremos aquí una clase de funciones denominadas **transformaciones lineales**, que transforman un vector $\mathbf{v}$ de un espacio vectorial V en otro vector $\mathbf{w}$ de un espacio vectorial W, cumpliendo con ciertas condiciones. Es decir, son casos especiales de funciones $F : V \to W$ entre dos espacios vectoriales. Se las puede puede pensar como *funciones de "una sola variable"*, donde el argumento de la función es un vector del espacio vectorial V, y el "valor" de la función es un vector del espacio vectorial W.

Si $V = W$, de modo que transforma un vector $\boldsymbol{v}$ de V en otro vector $\boldsymbol{w}$ del **mismo espacio** V, la transformación lineal se denomina usualmente **operador lineal**.

Usaremos la notación $L : V \longrightarrow W$ para describir una transformación lineal:

$$L(\mathbf{v}) = \mathbf{w}, \qquad \mathbf{v} \in V, \ \ \mathbf{w} \in W$$

Veremos que una transformación lineal L de un espacio vectorial n-dimensional V en otro espacio vectorial m-dimensional W podrá representarse por una matriz A de $m \times n$. Esto permitirá trabajar con la matriz A para discutir las características y propiedades de la transformación L, como así también, en ciertos casos, determinar las propiedades de la matriz A a partir de las propiedades de la transformación lineal que representa.

5.1.1. Definición general

Definición.
Una función $L : V \longrightarrow W$ de un espacio vectorial V en otro espacio vectorial W (ambos sobre el mismo conjunto de escalares) es una **transformación lineal** si satisface

$$L(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) = \alpha L(\mathbf{v}_1) + \beta L(\mathbf{v}_2)$$

para todo par de vectores $\mathbf{v}_1$ y $\mathbf{v}_2$ de V y todo par de escalares α, β. Es decir,

$$\begin{aligned} L(\mathbf{v}_1 + \mathbf{v}_2) &= L(\mathbf{v}_1) + L(\mathbf{v}_2) \quad \text{para todo } \mathbf{v}_1 \text{ y } \mathbf{v}_2 \text{ en } V \\ L(\alpha\mathbf{v}) &= \alpha L(\mathbf{v}) \quad \text{para todo } \mathbf{v} \text{ en } V \text{ y } \ \alpha \text{ escalar} \end{aligned}$$

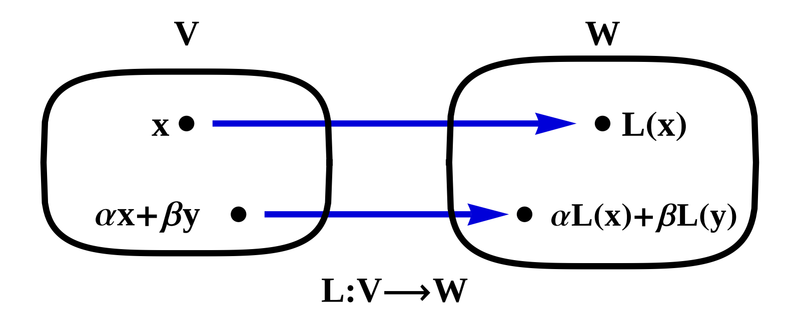


Figura 5.1: Esquema de una transformación lineal L entre espacios vectoriales.

En particular, para $\alpha = 0$ la última ecuación implica que **toda** transformación lineal $L : V \longrightarrow W$ satisface

$$L(\mathbf{0}_V) = \mathbf{0}_W$$

con $\mathbf{0}_V$ y $\mathbf{0}_W$ los vectores nulos de V y W, ya que $L(\mathbf{0}_V) = L(0\mathbf{v}) = 0L(\mathbf{v}) = \mathbf{0}_W$.

5.1.2. Transformaciones geométricas en $\mathbb{R}^2$

Ejemplos 5.1.1: Transformaciones lineales de $\Re^2$ en $\Re^2$
Comenzaremos con algunos ejemplos básicos:

1. **Transformación de dilatación (escalamiento):**

   Si $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ es un vector de $\mathbb{R}^2$, definimos, como primer ejemplo,

   $$L(\mathbf{x}) = 3\mathbf{x} = \begin{pmatrix} 3x_1 \\ 3x_2 \end{pmatrix}$$

   L es una transformación lineal, ya que

   $$\begin{aligned} L(\alpha\mathbf{x}) &= 3(\alpha\mathbf{x}) = \alpha(3\mathbf{x}) = \alpha L(\mathbf{x}) \\ L(\mathbf{x}+\mathbf{y}) &= 3(\mathbf{x}+\mathbf{y}) = 3\mathbf{x} + 3\mathbf{y} = L(\mathbf{x}) + L(\mathbf{y}) \end{aligned}$$

   verificándose que $L(\mathbf{0}) = \mathbf{0}$. Geométricamente, L tiene el efecto de "dilatar" el vector $\mathbf{x}$, multiplicando su longitud por un factor 3 y conservando su dirección y sentido:

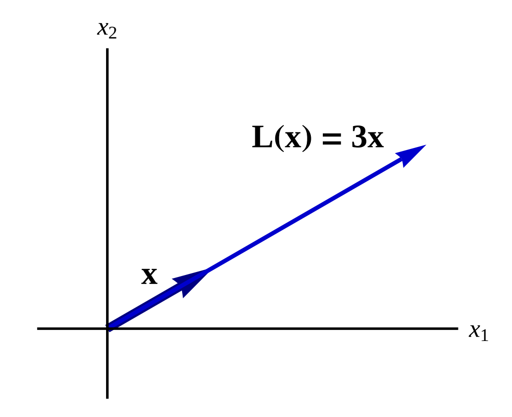


Figura 5.2: Dilatación (Escalamiento).

   Podemos expresar $L(\mathbf{x})$ en forma matricial como (¡verificar!)

   $$L(\mathbf{x}) = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

2. **Proyección ortogonal sobre el eje x_1:**

   Definimos ahora

   $$L(\mathbf{x}) = x_1\mathbf{e}_1 = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$$

Si $\mathbf{x}=\begin{pmatrix}x_1\\x_2\end{pmatrix}$, $\mathbf{y}=\begin{pmatrix}y_1\\y_2\end{pmatrix}$, entonces $\alpha\mathbf{x}+\beta\mathbf{y}=\begin{pmatrix}\alpha x_1+\beta y_1\\\alpha x_2+\beta y_2\end{pmatrix}$. Luego

$$\begin{aligned}L(\alpha\mathbf{x}+\beta\mathbf{y}) &= (\alpha x_1+\beta y_1)\,\mathbf{e}_1=\alpha(x_1\mathbf{e}_1)+\beta(y_1\mathbf{e}_1)\\&=\alpha L(\mathbf{x})+\beta L(\mathbf{y})\end{aligned}$$

lo que prueba que L es una transformación lineal.

Geométricamente, $L(\mathbf{x})$ es la proyección del vector $\mathbf{x}$ sobre el eje x_1

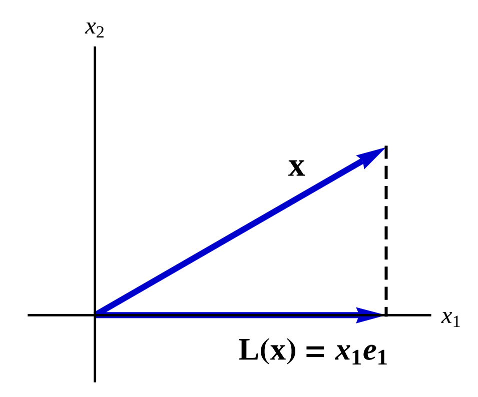


Figura 5.3: Proyección ortogonal sobre el eje x.

Podemos expresar $L(\mathbf{x})$ en forma matricial como (¡verificar!)

$$L(\mathbf{x})=\begin{pmatrix}1&0\\0&0\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}$$

3. **Reflexión respecto del eje x_1:**

Definimos

$$L(\mathbf{x})=x_1\mathbf{e}_1-x_2\mathbf{e}_2=\begin{pmatrix}x_1\\-x_2\end{pmatrix}$$

Esta tranformación satisface

$$\begin{aligned}L(\alpha\mathbf{x}+\beta\mathbf{y}) &= \begin{pmatrix}\alpha x_1+\beta y_1\\-(\alpha x_2+\beta y_2)\end{pmatrix}\\&=\alpha\begin{pmatrix}x_1\\-x_2\end{pmatrix}+\beta\begin{pmatrix}y_1\\-y_2\end{pmatrix}\\&=\alpha L(\mathbf{x})+\beta L(\mathbf{y})\end{aligned}$$

$\Rightarrow$ L es un operador lineal.

Geométricamente, $L(\mathbf{x})$ es la reflexión del vector $\mathbf{x}$ respecto (o a través) del eje x_1.

Podemos expresar $L(\mathbf{x})$ en forma matricial como (verificar)

$$L(\mathbf{x})=\begin{pmatrix}1&0\\0&-1\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}$$

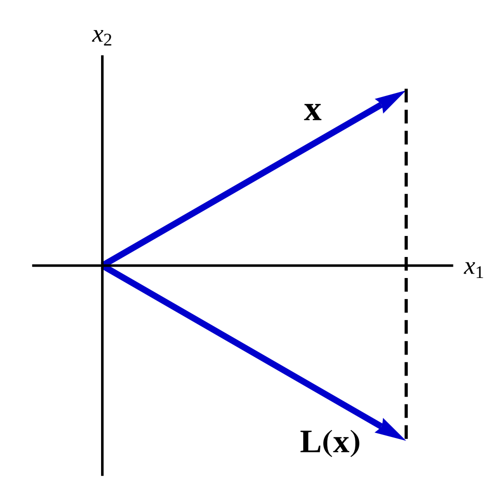


Figura 5.4: Reflexión respecto del eje x.

4. **Rotación de ángulo $\pi/2$ antihorario:**

   Si $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ definimos

$$L(\mathbf{x}) = -x_2\mathbf{e}_1 + x_1\mathbf{e}_2 = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$$

   Vemos que cumple

$$\begin{aligned} L(\alpha\mathbf{x} + \beta\mathbf{y}) &= \begin{pmatrix} -(\alpha x_2 + \beta y_2) \\ \alpha x_1 + \beta y_1 \end{pmatrix} = \alpha \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} + \beta \begin{pmatrix} -y_2 \\ y_1 \end{pmatrix} \\ &= \alpha L(\mathbf{x}) + \beta L(\mathbf{y}) \end{aligned}$$

   $\Rightarrow L$ es una transformación lineal.

   Geométricamente, $L(\mathbf{x})$ representa la rotación de ángulo $\theta = \pi/2$ (en sentido antihorario) del vector $\mathbf{x}$:

   Podemos expresar $L(\mathbf{x})$ en forma matricial como (verificar)

$$L(\mathbf{x}) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

5. **Transformación de escalamiento general**

   En general, el operador L_c definido por

$$L_c(\mathbf{x}) = c\mathbf{x} = \begin{pmatrix} cx_1 \\ cx_2 \end{pmatrix}$$

   con c un escalar fijo, es una transformación lineal, como podrá el lector probar fácilmente.

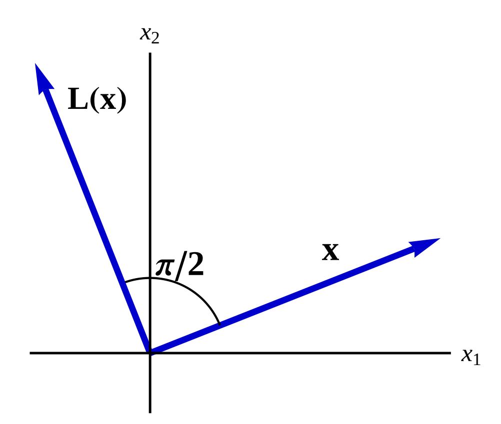


Figura 5.5: Rotación de ángulo $\pi/2$ antihoraria en el plano.

Si $c > 0$, L_c tiene el efecto de multiplicar la longitud del vector $\mathbf{x}$ por el factor de escala c, dilatando el vector un factor c si $c > 1$ y contrayendo el vector un factor c si $0 < c < 1$, pero siempre conservando su dirección y sentido.

Si $c = 1$, la transformación resultante

$$L_1(\mathbf{x}) = \mathbf{x}$$

se denomina **operador identidad** y se la denota como I: $I(\mathbf{x}) = \mathbf{x}\ \forall\ \mathbf{x} \in V$.
I no modifica ningún vector.

Si $c = 0$, la transformación resultante

$$L_0(\mathbf{x}) = \mathbf{0}$$

se denomina **operador nulo**. Envía a todos los vectores de V al vector nulo $\mathbf{0} \equiv \mathbf{0}_V$.

Si $c < 0$, L_c tendrá el efecto de invertir el sentido del vector, dilatándolo si $c < -1$, contrayéndolo si $-1 < c < 0$ y conservando su longitud si $c = -1$, en cuyo caso coincide con el operador de inversión.

Podemos expresar $L_c(\mathbf{x})$ en forma matricial como

$$L(\mathbf{x}) = \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

6. **Inversión:**

   Corresponde a

$$L(\mathbf{x}) = -\mathbf{x} = \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix}$$

   La linealidad de L es inmediata:

$$\begin{aligned} L(\alpha\mathbf{x} + \beta\mathbf{y}) &= \begin{pmatrix} -(\alpha x_1 + \beta y_1) \\ -(\alpha x_2 + \beta y_2) \end{pmatrix} = \alpha \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix} + \beta \begin{pmatrix} -y_1 \\ -y_2 \end{pmatrix} \\ &= \alpha L(\mathbf{x}) + \beta L(\mathbf{y}) \end{aligned}$$

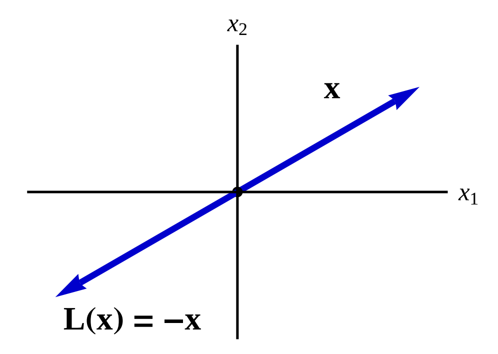


Figura 5.6: Inversión.

Geométricamente, $L(\mathbf{x})$ es el vector opuesto a $\mathbf{x}$.

Podemos expresar $L(\mathbf{x})$ en forma matricial como

$$L(\mathbf{x}) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Observar que el operador de inversión puede obtenerse como caso particular de otras transformaciones. Por ejemplo,

1. La transformación de escala con $c = -1$
2. Una rotación de ángulo π (en sentido anti-horario o sentido horario)

Esta última puede también lograrse mediante dos rotaciones sucesivas de ángulo $\pi/2$ (por ejemplo, ambas en sentido antihorario): Si R rota al vector $\mathbf{x}$ en $\pi/2$ antihorario, entonces

$$R(R(\mathbf{x})) = R\left(\begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}\right) = \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix} = -\mathbf{x}$$

Si definimos el cuadrado L^2 de un operador L (transformación de V en V) mediante

$$L^2(\mathbf{x}) \equiv L(L(\mathbf{x}))$$

entonces el operador de inversión L puede expresarse en términos del operador de rotación previo como

$$L = R^2$$

Ejemplos de transformaciones no lineales:

1. Si

$$F(\mathbf{x}) = -x_1 x_2 \mathbf{e}_1 + x_2 \mathbf{e}_2 = \begin{pmatrix} -x_1 x_2 \\ x_2 \end{pmatrix}$$

obtenemos

$$\begin{aligned} F(\alpha \mathbf{x}) &= -(\alpha x_1)(\alpha x_2)\mathbf{e}_1 + (\alpha x_2)\mathbf{e}_2 = \alpha \begin{pmatrix} \alpha x_1 x_2 \\ x_2 \end{pmatrix} \\ &\neq \alpha F(\mathbf{x}) \quad (\text{excepto para } \alpha = 0 \text{ o } 1 \text{ o } x_1 x_2 = 0) \end{aligned}$$

$\Rightarrow F$ no es una transformación lineal.

2. **Traslación:**

   La traslación $T_{\mathbf{a}}$ suma a todo vectir $\mathbf{x}$ un vector fijo $\mathbf{a}$:

$$T_{\mathbf{a}}(\mathbf{x}) = \mathbf{x} + \mathbf{a}$$

   Si $\mathbf{a} \neq \mathbf{0}$, $T_{\mathbf{a}}$ no es una transformación lineal, ya que por ejemplo, $T_{\mathbf{a}}(\mathbf{0}) = \mathbf{a} \neq \mathbf{0}$ y $T_{\mathbf{a}}(\mathbf{x}_1 + \mathbf{x}_2) = \mathbf{a} + \mathbf{x}_1 + \mathbf{x}_2 \neq T_{\mathbf{a}}(\mathbf{x}_1) + T_{\mathbf{a}}(\mathbf{x}_2)$.

   Así, no podrá representarse directamente mediante una matriz aplicada a $\mathbf{x}$.

Problemas 5.1.1

1. (i) Definir la proyección ortogonal en $\mathbb{R}^3$ sobre el plano-xy y mostrar que es una transformación lineal. Graficar.
   (ii) Definir la proyección ortogonal en $\mathbb{R}^3$ sobre el eje x y mostrar que es una transformación lineal. Graficar.
   (iii) ¿Qué es la proyección al origen? ¿ Puede considerarse una transformación lineal?

2. Considerar la transformación $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ dada por

$$L\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x/2 \\ y/3 \end{pmatrix}$$

   i) Verificar que es lineal. Expresarla en la forma matricial $L(\mathbf{x}) = A\mathbf{x}$.
   ii) Hallar las imágenes $L(\boldsymbol{v})$ de los vectores $\binom{1}{0}$, $\binom{0}{1}$, $\binom{1}{1}$
   iii) Dar una interpretación geométrica de L.
   iv) La imagen por L de un conjunto de vectores C se define como
   $L(C) = \{L(\mathbf{v}),\ \mathbf{v} \in C\}$. Probar que la imagen $L(C)$ bajo esta aplicación de la elipse

$$C = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \mid (x^2/4) + (y^2/9) = 1 \right\}$$

   es una circunferencia de radio 1.

Ejemplos de Transformaciones de $\mathbb{R}^n$ en $\mathbb{R}^m$

1. Sea $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ y $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^1$, definida por

$$L(\mathbf{x}) = x_1 + x_2$$

   L es una transformación lineal, ya que

$$\begin{aligned} L(\alpha\mathbf{x} + \beta\mathbf{y}) &= (\alpha x_1 + \beta y_1) + (\alpha x_2 + \beta y_2) \\ &= \alpha(x_1 + x_2) + \beta(y_1 + y_2) = \alpha L(\mathbf{x}) + \beta L(\mathbf{y}) \end{aligned}$$

   L asocia a cada vector $\mathbf{x} \in \mathbb{R}^2$ un escalar dado por $x_1 + x_2$. Puede ser expresada en forma matricial como $L(\mathbf{x}) = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

2. Sea $L:\mathbb{R}^2 \longrightarrow \mathbb{R}^3$ definida por $L(\mathbf{x}) = \begin{pmatrix} x_2 \\ x_1 \\ x_1+x_2 \end{pmatrix}$

   Se verifica fácilmente que L es lineal (¡probar!) y que puede ser escrita también como

$$L(\mathbf{x}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

5.1.3. Otros ejemplos

1. Dado un espacio vectorial V arbitrario, el operador **identidad** $I:V \longrightarrow V$ se define por

$$I(\mathbf{v}) = \mathbf{v} \quad \text{para todo } \mathbf{v} \in V$$

   Es, obviamente, una transformación lineal (¡verificar!).
   Notar que no existe $I: V \dashrightarrow W$ si $W \neq V$, aun si V y W tienen la misma dimensión.

2. La transformación nula $0:V \longrightarrow W$ se define por

$$0(\mathbf{v}) = \mathbf{0}_W \quad \text{para todo } \mathbf{v} \in V$$

   Es, obviamente, una transformación lineal (¡verificar!), que generaliza el operador nulo L_0 visto anteriormente.

3. **Transformación definida por una matriz** A.
   Dada una matriz A de $m \times n$, se puede definir una transformación lineal asociada $L:\mathbb{R}^n \longrightarrow \mathbb{R}^m$ dada por

$$L(\mathbf{x}) = A\,\mathbf{x}$$

   Es fácil ver que L cumple las propiedades de linealidad:

$$\begin{aligned} L(\alpha\mathbf{x}+\beta\mathbf{y}) &= A(\alpha\mathbf{x}+\beta\mathbf{y}) \\ &= \alpha A\mathbf{x} + \beta A\mathbf{y} \\ &= \alpha L(\mathbf{x}) + \beta L(\mathbf{y}) \end{aligned}$$

   Por lo tanto, cualquier matriz A de $m \times n$ puede verse como asociada a una transformación lineal $L:\mathbb{R}^n \longrightarrow \mathbb{R}^m$. Más aun, veremos luego que **toda** transformación lineal $L:\mathbb{R}^n \longrightarrow \mathbb{R}^m$ es de la forma anterior (para alguna matriz A de $m \times n$).

4. Sea $L: C_{[a,b]} \longrightarrow \mathbb{R}^1$ definida por

$$L(f) = \int_a^b f(x)\,dx$$

L es oviamente una transformación lineal, ya que si f y g son dos vectores cualesquiera de $C_{[a,b]}$, entonces

$$\begin{aligned} L(\alpha f+\beta g) &= \int_a^b (\alpha f+\beta g)(x)\,dx \\ &= \alpha\int_a^b f(x)\,dx+\beta\int_a^b g(x)\,dx \\ &= \alpha L(f)+\beta L(g) \end{aligned}$$

A diferencia de las anteriores, esta transformación lineal, cuyo dominio es un espacio vectorial de dimensión infinita, no puede representarse mediante una matriz.

5. Sea $D : C^\infty \longrightarrow C^\infty$ el operador derivada en el espacio C^∞ de funciones reales $f : \mathbb{R} \longrightarrow \mathbb{R}$ derivables a todo orden, definida por

$$D(f) = f'$$

es decir, $D(f)(x) = f'(x)$. Se la suele denotar directamente como $D = \frac{d}{dx}$.
D es obviamente un operador lineal, ya que si f y $g \in C^\infty$,

$$D(\alpha f+\beta q) = (\alpha f+\beta g)' = \alpha f'+\beta g' = \alpha D(f)+\beta D(g)$$

Nótese que D^2 es el operador derivada segunda $\frac{d^2}{dx^2}$:

$$D^2(f) = D(D(f)) = D(f') = f''$$

Dado que C^∞ tiene dimensión infinita, D no puede representarse mediante una matriz (pero si se restringe el dominio a un subespacio de C^∞ de dimensión finita, tal como el espacio P_n de polinomios de grado $\leq n$, D sí podrá representarse mediante una matriz, como veremos luego).

Importante: Si $L : V \to W$ es una transformación lineal, se cumplen siempre las siguientes reglas o propiedades:

1. $L(\mathbf{0}_V) = \mathbf{0}_W$
2. Si $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$, entonces

$$L(\alpha_1\mathbf{v}_1+\cdots+\alpha_n\mathbf{v}_n) = \alpha_1 L(\mathbf{v}_1)+\cdots+\alpha_n L(\mathbf{v}_n)$$

3. $$L(-\mathbf{v}) = -L(\mathbf{v}) \quad \forall \;\; \mathbf{v} \in V$$

Se dejan las demostraciones para el lector.

Problema 5.1.2

1. Sea $L : V \longrightarrow W$ una transformación lineal y sean $\boldsymbol{w}_1 = L(\boldsymbol{v}_1), \ldots, \boldsymbol{w}_k = L(\boldsymbol{v}_k)$ las imágenes de k vectores $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ de V.
   a) Mostrar que si el conjunto de los vectores $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ es linealmente dependiente $\Rightarrow \{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k\}$ es linealmente dependiente.
   b) Mostrar que si $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ es linealmente independiente $\Rightarrow$ el conjunto $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k\}$ no es necesariamente independiente. Dar un ejemplo (considere proyecciones ortogonales sobre un cierto eje o la transformación nula).

5.2. Imagen y núcleo de una transformación lineal

Sea $L : V \longrightarrow W$ una transformación lineal

1. **Núcleo** de L: es el conjunto de vectores $\mathbf{v}$ de V que son transformados o enviados al vector nulo $\mathbf{0}_W$ de W. Es decir,

$$\mathrm{Nu}\,(L) = \{\mathbf{v} \in V : L\,(\mathbf{v}) = \mathbf{0}_W\}$$

2. **Imagen** de un subespacio S de V: es el conjunto de vectores $\mathbf{w}$ de W que son imagen por L de vectores $\mathbf{v}$ de S, es decir,

$$L\,(S) = \{\mathbf{w} \in W : \mathbf{w} = L\,(\mathbf{v})\,\text{para algún}\ \mathbf{v} \in S\} = \{L(\mathbf{v}),\ \mathbf{v} \in S\}$$

3. **Imagen de** L: Es la imagen $L\,(V)$ de todo el espacio vectorial V:

$$\mathrm{Im}(L) = L(V) = \{L(\mathbf{v}),\ \mathbf{v} \in V\} \subset W$$

Notar que la *imagen* por L del $\mathrm{Nu}(L)$ es el vector nulo $\mathbf{0}_W$ de W: $L(\mathrm{Nu}(L)) = \{\mathbf{0}_W\}$.

Cada uno de estos conjuntos de vectores es un subespacio en los respectivos espacios vectoriales:

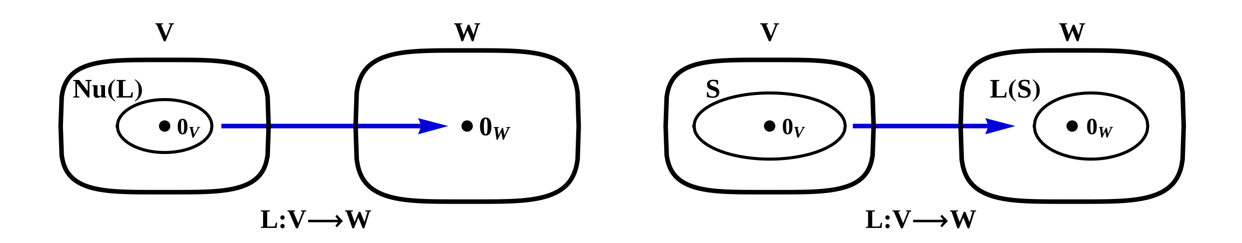


Figura 5.7: Esquema de Núcleo e Imagen de una transformación lineal.

Teorema
Si $L : V \longrightarrow W$ es una transformación lineal, entonces

1. $\mathrm{Nu}\,(L)$ es un subespacio de V
2. Si S es un subespacio de V, $L\,(S)$ es un subespacio de W. Esto implica en particular que la imagen $\mathrm{Im}(L) = L(V)$ es un subespacio de W.
3. Si V es de dimensión finita, la suma de la dimensión de la imagen $\mathrm{Im}(L)$ y la dimensión del núcleo $\mathrm{Nu}(L)$ es la dimensión del espacio V:

$$\dim \mathrm{Im}\,(L) + \dim \mathrm{Nu}\,(L) = \dim V$$

Demostración de 1. En primer lugar, $L(\mathbf{0}_V) = \mathbf{0}_W$, por lo que $\mathbf{0}_V \in \mathrm{Nu}(L)$. Además, si $\mathbf{v}_1$ y $\mathbf{v}_2 \in \mathrm{Nu}\,(L)$,

$$\begin{aligned} L\,(\mathbf{v}_1 + \mathbf{v}_2) &= L\,(\mathbf{v}_1) + L\,(\mathbf{v}_2) = \mathbf{0}_W + \mathbf{0}_W = \mathbf{0}_W \\ L\,(\alpha\mathbf{v}_1) &= \alpha L\,(\mathbf{v}_1) = \mathbf{0}_W \end{aligned}$$

por lo que $\mathbf{v}_1 + \mathbf{v}_2$ y $\alpha\mathbf{v}_1 \in \mathrm{Nu}\,(L)$. El núcleo es pues cerrado por la suma y multiplicación por un escalar, siendo entonces un subespacio.

Demostración de 2. $L(S)$ contiene al $\mathbf{0}_W$ pues $L(\mathbf{0}_V) = \mathbf{0}_W$. Además, si $\mathbf{w}_1$ y $\mathbf{w}_2$ son vectores en $L\,(S)$, existen $\mathbf{v}_1$ y $\mathbf{v}_2$ en S tales que $\mathbf{w}_1 = L(\mathbf{v}_1)$, $\mathbf{w}_2 = L(\mathbf{v}_2)$. Entonces

$$\begin{aligned} \alpha\mathbf{w}_1 &= \alpha L\,(\mathbf{v}_1) = L\,(\alpha\mathbf{v}_1) \quad \text{para } \mathbf{v}_1 \in S \\ \mathbf{w}_1 + \mathbf{w}_2 &= L\,(\mathbf{v}_1) + L\,(\mathbf{v}_2) = L\,(\mathbf{v}_1 + \mathbf{v}_2) \qquad \text{para } \mathbf{v}_1 \text{ y } \mathbf{v}_2 \,\in\, S \end{aligned}$$

Como S es un subespacio, ambos $\alpha\mathbf{v}_1$ y $\mathbf{v}_1 + \mathbf{v}_2$ pertenecen también a S, por lo que $\alpha\mathbf{w}_1 \in L(S)$ y $\mathbf{w}_1 + \mathbf{w}_2 \in L(S)$. Luego, $L\,(S)$ es cerrado por la suma y multiplicación por un escalar, siendo entonces un subespacio.

Demostración de 3. Partiendo de una base $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ de V tal que $\{\boldsymbol{v}_{k+1}, \ldots, \mathbf{v}_n\}$ es una base de $\mathrm{Nu}(L)$ ($L(\mathbf{v}_{k+1}) = \ldots = L(\mathbf{v}_n) = \mathbf{0}_W$), todo $\mathbf{v} \in V$ puede escribirse como $\mathbf{v} = \alpha_1\mathbf{v}_1 + \ldots + \alpha_k\mathbf{v}_k + \alpha_{k+1}\mathbf{v}_{k+1} + \ldots + \alpha_n\mathbf{v}_n$. Entonces,

$$\begin{aligned} L(\mathbf{v}) &= L(\alpha_1\mathbf{v}_1 + \ldots + \alpha_k\mathbf{v}_k + \alpha_{k+1}\mathbf{v}_{k+1} + \ldots + \alpha_n\mathbf{v}_n) \\ &= \alpha_1 L(\mathbf{v}_1) + \ldots + \alpha_k L(\mathbf{v}_k) + \alpha_{k+1} L(\mathbf{v}_{k+1}) + \ldots + \alpha_n L(\mathbf{v}_n) \\ &= \alpha_1 L(\mathbf{v}_1) + \ldots + \alpha_k L(\mathbf{v}_k) \end{aligned}$$

por lo que $\{L(\mathbf{v}_1), \ldots, L(\mathbf{v}_k)\}$ **genera** $\mathrm{Im}(L)$. Además, $\{L(\mathbf{v}_1), \ldots, L(\mathbf{v}_k)\}$ es linealmente independiente, pues si $\mathbf{0}_W = \alpha_1 L(\mathbf{v}_1) + \ldots + \alpha_k L(\mathbf{v}_k) = L(\alpha_1\mathbf{v}_1 + \ldots + \alpha_k\mathbf{v}_k) \Rightarrow$ $\alpha_1\mathbf{v}_1 + \ldots + \alpha_k\mathbf{v}_k \in \mathrm{Nu}(L)$ y entonces $\alpha_1\mathbf{v}_1 + \ldots + \alpha_k\mathbf{v}_k = \beta_1\mathbf{v}_{k+1} + \ldots + \beta_n\mathbf{v}_n$. Pero por ser los $\mathbf{v}_i$ linealmente independientes, $\alpha_1 = \ldots = \alpha_k = \beta_{k+1} = \ldots = \beta_n = 0$. Por lo tanto,

$$\dim \mathrm{Im}(L) + \dim \mathrm{Nu}(L) = k + (n - k) = n = \dim V$$

Ejemplos 5.2

1. Sea $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida por

$$L(\mathbf{x}) = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$$

(operador de proyección). Es obvio que $L(\mathbf{x}) = 0$ si y sólo si $x_1 = 0$, es decir, $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathrm{Nu}(L)$ si y sólo si $x_1 = 0$. Por lo tanto, $\mathrm{Nu}(L)$ es el subespacio 1-dimensional de $\mathbb{R}^2$ generado por el vector $\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, es decir, es el eje x_2, como es obvio geométricamente (¡graficar!):

$$\mathrm{Nu}(L) = \left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle$$

Por otra parte, dado que $L(\mathbf{x}) = x_1\mathbf{e}_1$, la imagen $\mathrm{Im}(L) = L(V)$ es el conjunto de vectores proporcionales a $\mathbf{e}_1$, es decir, el subespacio 1-dimensional de $\mathbb{R}^2$ generado por el vector $\mathbf{e}_1$, que geométricamente es el eje x_1:

$$\mathrm{Im}(L) = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle$$

Se verifica que $\dim \mathrm{Im}(L) + \dim \mathrm{Nu}(L) = 1 + 1 = 2 = \dim V$ $(V = \mathbb{R}^2)$.
Nótese que $L(\mathbf{x}) = A\mathbf{x}$, con $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, y que $\mathrm{Nu}(L)$ coincide con el espacio nulo de $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, mientras que $\mathrm{Im}(L)$ coincide con el espacio columna de A.

2. Sea $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ definida por

$$L(\mathbf{x}) = \begin{pmatrix} x_1 + x_2 \\ x_2 + x_3 \end{pmatrix}$$

Si $\mathbf{x} \in \mathrm{Nu}(L)$, entonces $\left\{ \begin{array}{rcl} x_1 + x_2 & = & 0 \\ x_2 + x_3 & = & 0 \end{array} \right.$. Resolviendo el sistema, si la variable independiente es $x_3 = t$, se tiene $x_2 = -x_3$, $x_1 = x_3$, es decir, $\mathbf{x} = \begin{pmatrix} t \\ -t \\ t \end{pmatrix} = t \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$:

$$\mathrm{Nu}(L) = \left\langle \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\rangle$$

Por otro lado, vemos que

$$L(\mathbf{x}) = x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

con x_1, x_2, x_3 arbitrarios, por lo que la imagen será $\mathbb{R}^2$:

$$\mathrm{Im}(L) = \left\langle \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix} \right\rangle = \left\langle \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix} \right\rangle = \mathbb{R}^2$$

Se verifica que $\dim \mathrm{Im}(L) + \dim \mathrm{Nu}(L) = 2 + 1 = 3 = \dim V$ ($V = \mathbb{R}^3$). Nótese también que $L(\mathbf{x}) = A\mathbf{x}$, con $A = \begin{pmatrix}1 & 1 & 0\\0 & 1 & 1\end{pmatrix}$, coincidiendo $\mathrm{Nu}(L)$ con el espacio nulo de A y $\mathrm{Im}(L)$ con el espacio columna de A (ver problema 5.2.7).

Problemas 5.2

1. Verificar que las siguientes transformaciones $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ son lineales, y determinar $\mathrm{Nu}(L)$, la imagen $\mathrm{Im}(L)$ y sus dimensiones, junto con una base de los mismos:
   (a) $L\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}x\\x+y+z\end{pmatrix}$ **(b)** $L\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}2x+y\\-4x-2y\end{pmatrix}$

2. Idem 1. para las siguientes transformaciones $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$:
   (a) $L\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}y\\-x\end{pmatrix}$ **(b)** $L\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}0\\y\end{pmatrix}$
   Interprételas geométricamente.

3. **Importante:** Sea $L : \mathbb{R}^n \to \mathbb{R}^m$ la transformación lineal definida por una matriz A de $m \times n$:
   $$L(\mathbf{x}) = A\mathbf{x}$$
   Probar que:
   a) El núcleo $\mathrm{Nu}(L)$ es el **espacio nulo** de la matriz A.
   b) La imagen $\mathrm{Im}(L)$ es el **espacio generado por las columnas** de la matriz A (o sea, el espacio columna de la matriz).
   c) La igualdad dim $\mathrm{Im}(L)$ +dim $\mathrm{Nu}(L)$ = dim V es equivalente a
   $$\text{rango }(A) + \text{nulidad }(A) = n$$
   d) Verifique estos resultados para las transformaciones del ejercicio 2., escribiéndolas en la forma $L(\mathbf{x}) = A\mathbf{x}$.

4. Determinar si son lineales las siguientes transformaciones $L : \mathbb{R}^{2\times 2} \longrightarrow \mathbb{R}$. En caso de que lo sean halle su núcleo e imagen.
   (a) $L(\begin{pmatrix}a & b\\c & d\end{pmatrix}) = a + d$ (Traza de la matriz)
   (b) $L(\begin{pmatrix}a & b\\c & d\end{pmatrix}) = ad - bc$ (Determinante de la matriz)

5. a) Determine si la traza de una matriz cuadrada A de $n \times n$, definida por
   $$Tr(A) = \sum_{i=1}^{n} a_{ii} = a_{11} + \ldots + a_{nn}$$

es una transformación lineal $T : \mathbb{R}^{n\times n} \longrightarrow \mathbb{R}$.
b) Indique si el determinante $\det(A)$ de una matriz cuadrada A de $n \times n$, es una transformación lineal $\det : \mathbb{R}^{n\times n} \longrightarrow \mathbb{R}$.

6. Halle el núcleo e imagen del operador identidad $I : V \to V$, y del operador nulo $0 : V \to V$.

7. Mostrar que una función $f : \mathbb{R} \to \mathbb{R}$ cuya gráfica es una recta no es necesariamente una transformación lineal $L : \mathbb{R} \to \mathbb{R}$ (considere por ejemplo la recta $y = mx + 1$).

8. Mostrar que $\forall\, k \geq 1$, la derivada k-ésima $D^k = d^k/dt^k$ en el espacio P de polinomios (de grado arbitrario) es una transformación lineal. ¿Cuál es su núcleo e imagen?

9. **Importante: Propiedades geométricas de una transformación lineal**.

   a) Pobar que toda transformación lineal $L : \mathbb{R}^2 \to \mathbb{R}^2$ con $\mathrm{Nu}(L) = \{\mathbf{0}_V\}$ transforma rectas $R = \{\mathbf{r}_0 + t\mathbf{n},\ t \in \mathbb{R}\}$, $\mathbf{n} \neq \mathbf{0}$, en rectas, y segmentos $Q = \{\mathbf{r}_0 + t\mathbf{n},\ t \in [t_0, t_1]\}$) en segmentos. ¿Qué puede suceder si $\dim \mathrm{Nu}(L) \geq 1$? ¿Siguen siendo válidos estos resultados en $\mathbb{R}^3$? ¿ y en $\mathbb{R}^n$?

   b) Probar que toda transformación lineal $L : \mathbb{R}^2 \to \mathbb{R}^2$ con $\mathrm{Nu}(L) = \{\mathbf{0}_V\}$ transforma rectas y segmentos paralelos (caracterizados por un mismo vector $\mathbf{n} \neq \mathbf{0}$ pero distintos $\mathbf{r}_0$) en rectas y segmentos paralelos. Generalizar a $\mathbb{R}^3$ y $\mathbb{R}^n$. ¿Puede extenderse el resultado a planos paralelos en $\mathbb{R}^3$?

   c) Si $L(\mathbf{x}) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$, determine la imagen por L de las rectas paralelas $y = 2x$ y $y = 2x + 1$. Grafique e interprete el resultado.

   d) Dados $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$, i) probar que el segmento de recta que los conecta es el conjunto $Q = \{t\mathbf{v} + (1-t)\mathbf{u},\ t \in [0,1]\}$. Verifíquelo para $\mathbf{u} = (1,0)$, $\mathbf{v} = (0,1)$. ii) Mostrar que su imagen $L(Q)$ bajo una transformación lineal $L : \mathbb{R}^2 \to \mathbb{R}^2$ es el segmento de recta entre $L(\mathbf{u})$ y $L(\mathbf{v})$. Generalizar a $\mathbb{R}^3$ y $\mathbb{R}^n$.
   iii) Si $L(\mathbf{x}) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$, determine la imagen por L del segmento de recta entre $\mathbf{u} = (1,0)$ y $\mathbf{v} = (0,1)$. Graficar.

   e) Un subconjunto $C \subset \mathbb{R}^n$ es **convexo** si para cualquier par de puntos de C el segmento de recta que los conecta yace enteramente en C, es decir,

$$\mathbf{u} \in C, \mathbf{v} \in C \ \Rightarrow\ t\mathbf{v} + (1-t)\mathbf{u} \ \in\ C\ \forall\, t \in [0,1]$$

   Por ejemplo, todo subespacio de $\mathbb{R}^n$ es un conjunto convexo (¡probar!). Pero un conjunto convexo no necesariamente es un subespacio. Por ejemplo, en $\mathbb{R}^2$ un círculo "lleno" $C = \{(x,y) \in \mathbb{R}^2, x^2 + y^2 \leq 1\}$ es un conjunto convexo (¡probar!). En cambio, el círculo $\{(x,y) \in \mathbb{R}^2, x^2 + y^2 = 1\}$ no es un conjunto convexo.

   Probar que toda transformación lineal $L : \mathbb{R}^n \to \mathbb{R}^m$ transforma un conjunto convexo en un conjunto convexo. Dé un ejemplo.

5.3. Propiedades fundamentales

1. Si $L : V \longrightarrow W$ es una transformación lineal, con V un espacio vectorial de dimensión finita n y $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base arbitraria de V, entonces L queda completamente determinada por los n vectores $\{L(\mathbf{v}_1), \ldots, L(\mathbf{v}_n)\}$, es decir, por las n imágenes de los vectores de la base.

   En efecto, si $\mathbf{v} \in V$, entonces

   $$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n$$

   y por lo tanto

   $$L(\mathbf{v}) = L(\alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n) = \alpha_1 L(\mathbf{v}_1) + \ldots + \alpha_n L(\mathbf{v}_n)$$

   es decir, $L(\mathbf{v})$ es una combinación lineal de los n vectores $L(\mathbf{v}_1), \ldots, L(\mathbf{v}_n)$.
   La imagen $L(V)$ es entonces el espacio generado por estos n vectores:

   $$L(V) = \langle L(\mathbf{v}_1), \ldots, L(\mathbf{v}_n) \rangle$$

   Nótese, no obstante, que el conjunto $\{L(\mathbf{v}_1), \ldots, L(\mathbf{v}_n)\}$ puede ser linealmente dependiente (por ej., algunos vectores $L(\mathbf{v}_i)$ pueden ser nulos), en cuyo caso no será una base de $L(V)$.

2. Una transformación lineal $L : V \longrightarrow W$ es inyectiva (o monomorfismo) si
   $L(\mathbf{v}_1) \neq L(\mathbf{v}_2)\ \forall\ \boldsymbol{v}_1 \neq \boldsymbol{v}_2$. Es fácil ver que es inyectiva **si y sólo si** $\mathrm{Nu}(L) = \{\mathbf{0}_V\}$.

   En efecto, L inyectiva $\Rightarrow L(\mathbf{v}) \neq L(\mathbf{0_V}) = \mathbf{0}_W\ \forall\ \mathbf{v} \neq \mathbf{0}_V$, por lo que $\mathrm{Nu}(L) = \{\mathbf{0}_V\}$.
   Y si $\mathrm{Nu}(L) = \{\mathbf{0}_V\} \Rightarrow L(\mathbf{v}_1) - L(\mathbf{v}_2) = L(\mathbf{v}_1 - \mathbf{v}_2) \neq \mathbf{0}_W\ \forall\ \mathbf{v}_1 \neq \mathbf{v}_2$, por lo que L es inyectiva.

3. Si $\mathrm{Nu}(L) = \{\mathbf{0}_V\}$ y $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subset V$ es linealmente independiente, el conjunto $\{L(\mathbf{v}_1), \ldots, L(\mathbf{v}_k)\}$ es **linealmente independiente**. En otras palabras, si la transformación lineal L es inyectiva, **conserva la independencia lineal**.
   En efecto, si

   $$\mathbf{0}_W = \alpha_1 L(\mathbf{v}_1) + \ldots + \alpha_k L(\mathbf{v}_k) = L(\alpha_1 \mathbf{v}_1 + \ldots + \alpha_k \mathbf{v}_k)$$

   entonces $\alpha_1 \mathbf{v}_1 + \ldots + \alpha_k \mathbf{v}_k \in \mathrm{Nu}(L)$, por lo que $\alpha_1 \mathbf{v}_1 + \ldots + \alpha_k \mathbf{v}_k = \mathbf{0}_V$. Pero esto implica $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$ por ser $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ linealmente independiente. Por lo tanto, $\{L(\mathbf{v}_1), \ldots, L(\mathbf{v}_k)\}$ es **linealmente independiente**.

   En particular, si $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es una base de $V \Rightarrow \{L(\mathbf{v}_1), \ldots, L(\mathbf{v}_n)\}$ es una **base** de la imagen $\mathrm{Im}(L) = L(V)$, pues por lo anterior, es linealmente independiente y por la propiedad 1, **genera** la imagen. Por lo tanto, para $L : V \to W$ inyectiva, $\dim \mathrm{Im}(L) = n$ y $\dim \mathrm{Nu}(L) = 0$, verificándose que

   $$\dim \mathrm{Im}(L) + \dim \mathrm{Nu}(L) = n + 0 = n = \dim V$$

   Esto también implica que si $L : V \to W$ es inyectiva, necesariamente
   $\dim V \leq \dim W$ pues $\mathrm{Im}(V) \subset W$.

4. Si $L : V \longrightarrow W$ es una transformación lineal y $L(V) = W \Rightarrow L$ es sobreyectiva (epimorfismo). En este caso la imagen $L(V)$ es todo el codominio W, es decir, cubre todo W, por lo que el conjunto $\{L(\mathbf{v}_1), \ldots, L(\mathbf{v}_n)\}$ **genera** W.

   Además, en este caso se cumple que

$$\dim V = \dim \text{Im}(L) + \dim \text{Nu}(L) = \dim W + \dim \text{Nu}(L)$$

   por lo que necesariamente $\dim V \geq \dim W$.

5.3.1. Isomorfismos

1. Si $L : V \Rightarrow W$ es una transformación lineal **biyectiva**, es decir, que es a la vez inyectiva ($\text{Nu}(L) = \{\mathbf{0}_V\}$) y sobreyectiva ($L(V) = W$) $\Rightarrow$ se dice que L es un **isomorfismo**. Si $V = W$ al isomorfismo se lo denota **automorfismo**.

   Si L es un isomorfismo y dim $V = n$, con $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base de $V \Rightarrow$

$$\{L(\mathbf{v}_1), \ldots, L(\mathbf{v}_n)\}$$

   es una **base** de W, pues son linealmente independientes (por ser L inyectiva) y generan W (por ser L sobreyectiva). Un isomorfismo transforma entonces cualquier **base** de V en una **base** de W.

   Por lo tanto, V y W **deben tener la misma dimensión** n (cuando son de dimensión finita). Para un isomorfismo se verifica entonces que

$$\dim \text{Im}(L) + \dim \text{Nu}(L) = n + 0 = \dim V = \dim W$$

2. Una función $L : V \longrightarrow W$ tiene inversa $L^{-1} : W \longrightarrow V$ si y sólo si L es biyectiva. Por lo tanto, si $L : V \longrightarrow W$ es una transformación lineal, L tendrá inversa $L^{-1} : W \longrightarrow V$ **si y sólo si** L **es un isomorfismo:**

$$L(\mathbf{v}) = \mathbf{w} \Rightarrow \mathbf{v} = L^{-1}(\mathbf{w})$$

   cumpliéndose que

$$L(L^{-1}(\mathbf{w})) = L(\mathbf{v}) = \mathbf{w} \quad \forall\, \mathbf{w} \in W$$
$$L^{-1}(L(\mathbf{v})) = L^{-1}(\mathbf{w}) = \mathbf{v} \quad \forall\, \mathbf{v} \in V$$

   es decir,

$$LL^{-1} = I_W, \quad L^{-1}L = I_V$$

   donde I_W y I_V son los operadores identidad en W y V.
   La inversa $L^{-1} : W \longrightarrow V$ de un isomorfismo L **es también un isomorfismo** (¡probar!), es decir, una transformación lineal biyectiva.

3. Se dice que dos espacios vectoriales V, W son **isomorfos** si existe un isomorfismo $L : V \longrightarrow W$ entre ellos. Si V y W son de dimensión finita $\Rightarrow V$ y W son isomorfos **si y sólo si tienen la misma dimensión** (¡probar!).

Ejemplo 5.3 Sea $L : \mathbb{R}^2 \Rightarrow \mathbb{R}^2$ la transformación dada por

$$L(\mathbf{x}) = \begin{pmatrix} x_1 + x_2 \\ x_1 - x_2 \end{pmatrix}$$

Se verifica en primer lugar que L es lineal y que $\forall\ \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 \in \mathbb{R}^2$,

$$L(\mathbf{x}) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} x_2 \\ -x_2 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} = x_1 L(\mathbf{e}_1) + x_2 L(\mathbf{e}_2)$$

con $L(\mathbf{e}_1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $L(\mathbf{e}_2) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, por lo que basta conocer $L(\mathbf{e}_1)$ y $L(\mathbf{e}_2)$ para determinar $L(\mathbf{x})$ para cualquier $\mathbf{x} \in \mathbb{R}^2$.

Se comprueba también que $\mathrm{Nu}(L) = \mathbf{0}$, pues si $x_1 + x_2 = 0$ y $x_1 - x_2 = 0 \Rightarrow x_1 = x_2 = 0$ (¡verificar!). Por lo tanto L es **inyectiva**. Y finalmente, vemos que la imagen de L es

$$L(V) = \{x_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix},\ x_1, x_2 \in \mathbb{R}\} = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle = \mathbb{R}^2$$

por lo que L es también **sobreyectiva**. Este resultado puede también obtenerse directamente de

$$\dim \mathrm{Im}(L) = \dim V - \dim \mathrm{Nu}(L) = 2 - 0 = 2$$

L es entonces un **automorfismo**, es decir, un **isomorfismo** entre V y V, con $V = \mathbb{R}^2$.

L tendrá entonces inversa $L^{-1} : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, dada por (¡verificar!)

$$L^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} x_1 + x_2 \\ x_1 - x_2 \end{pmatrix}$$

Notemos finalmente que podemos expresar $L(\mathbf{x})$ como

$$L(\mathbf{x}) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

y $L^{-1}(\mathbf{x})$ como

$$L^{-1}(\mathbf{x}) = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

siendo la matriz que representa a L^{-1} la **inversa** de la matriz que representa a L.

Problemas 5.3

1. Si $L : V \longrightarrow V$ es un operador lineal y $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es una base de V, probar
   (a) Si $L(\mathbf{v}_i) = \mathbf{0}$ para cada elemento de la base entonces L es la transformación lineal nula.
   (b) Si $L(\mathbf{v}_i) = \mathbf{v}_i$ para cada vector de la base entonces L es la identidad.
   (c) Si existe un escalar r tal que $L(\mathbf{v}_i) = r\mathbf{v}_i$ para cada vector de la base entonces $L(\mathbf{v}) = r\mathbf{v}$ para todo los vectores en V.

2. Sea $L\colon V \to W$ una transformación lineal y supongamos que $L(\mathbf{v}_1) = \mathbf{w}_1, \ldots,$ $L(\mathbf{v}_k) = \mathbf{w}_k$ para vectores $\mathbf{v}_1, \ldots, \mathbf{v}_k$ de V.
   (a) Si $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ genera W, ¿debe $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ generar V? Pensar, por ejemplo, en transformaciones de $\Re^3$, sobre $\Re^2$.
   (b) Si $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ genera V, ¿debe $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ generar W? Pensar por ejemplo en $L : \Re^2 \longrightarrow \Re^3$.
   (c) Si ahora L es un isomorfismo (que implica dim $V =$ dim W en espacios de dimensión finita) ¿ cuál es la respuesta a (a) y (b)?

3. Si $L : V \to W$ es inyectiva, mostrar que la imagen $L(S)$ de un subespacio de dimensión m de V es un subespacio de dimensión m de W.

4. **Importante:** Sea $L : \mathbb{R}^n \to \mathbb{R}^m$ la transformación lineal definida por una matriz A de $m \times n$:
$$L(\mathbf{x}) = A\mathbf{x}$$
   probar que:
   a) L es inyectiva si y sólo si rango$(A) = n$. Esto implica $n \leq m$.
   b) L es sobreyectiva si y sólo si rango$(A) = m$. Esto implica $n \geq m$.
   c) L es biyectiva (isomorfismo) si y sólo si rango$(A) = m = n$, es decir, si y sólo si A es una matriz cuadrada no singular.
   d) Probar que en c), $L^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ está dada por
$$L^{-1}(\mathbf{x}) = A^{-1}(\mathbf{x})$$
   e) Discuta las implicancias de los resultados anteriores para el sistema de ecuaciones lineales (de $m \times n$)
$$L(\mathbf{x}) = \mathbf{b}$$
   En particular, muestre que
   i) Es compatible si y sólo si $\mathbf{b} \in$ Im(L).
   ii) Es compatible $\forall\ \boldsymbol{b} \in \mathbb{R}^m$ si y sólo si L es sobreyectiva
   iii) La solución, cuando existe, es única si y sólo si L es inyectiva
   iv) La solución existe y es única $\forall\ \boldsymbol{b} \in \mathbb{R}^m$ si y sólo si L es biyectiva (isomorfismo), es decir si y sólo si $m = n$ y A es no singular.

5. **Importante:** Sea V un espacio vectorial de dimensión n, con $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base (ordenada) de V, tal que si $\mathbf{v} \in V$,
$$\mathbf{v} = x_1\mathbf{v}_1 + \ldots + x_n\mathbf{v}_n$$
   Sea $L : V \longrightarrow \mathbb{R}^n$ la transformación lineal definida por
$$L(\mathbf{v}) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$
   Es decir, $L(\mathbf{v}) = [\mathbf{v}]_B$ es el vector columna de coordenadas de $\mathbf{v}$ en la base B.
   a) Mostrar que L está bien definida, es decir, que $[\mathbf{v}]_B$ existe y es único $\forall\ \mathbf{v} \in V$.

b) Mostrar que L es una transformación lineal.
c) Mostrar que L es un isomorfismo (es decir, $\mathrm{Nu}(L) = \{\mathbf{0}_V\}$, $\mathrm{Im}(L) = \mathbb{R}^n$).
Este resultado muestra en forma explícita que todo espacio V de dimensión n (con escalares reales) es isomorfo a $\mathbb{R}^n$, es decir, que existe una correspondencia biyectiva entre ambos.
d) ¿Cuál es la inversa L^{-1}?

6. Mostrar que dos espacios V, W son isomorfos si y sólo si tienen la misma dimensión.

7. Mostrar que la inversa de un isomorfismo $L : V \longrightarrow W$ es un isomorfismo, es decir, que L^{-1} es lineal y biyectiva.

5.4. Representación matricial de transformaciones lineales

Veremos ahora que cualquier transformación lineal L entre espacios vectoriales de dimensión finita V y W se puede representar mediante una matriz A, que dependerá de las bases que se consideren en V y W.

En primer lugar consideramos transformaciones de $\mathbb{R}^n$ en $\mathbb{R}^m$:

$$L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

Asumimos primero que la base en $V = \mathbb{R}^n$ es la base canónica $B_c = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$.
Dado $\mathbf{x} \in \mathbb{R}^n$, podemos representarlo como

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n$$

Como L es lineal,

$$L(\mathbf{x}) \quad = \quad x_1 L(\mathbf{e}_1) + \ldots + x_n L(\mathbf{e}_n)$$

Luego si para cada $\mathbf{e}_j$, $j = 1, \ldots, n$, se tiene

$$L(\mathbf{e}_j) = \mathbf{a}_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

entonces

$$\begin{aligned} L(\mathbf{x}) &= x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \ldots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \\ &= \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \end{aligned}$$

es decir,

$$L(\mathbf{x}) = A\mathbf{x}$$

con A la matriz de $\boldsymbol{m} \times \boldsymbol{n}$

$$A = (L(\mathbf{e}_1), \ldots, L(\mathbf{e}_n)) = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}$$

La matriz A se denomina **representación matricial** de L en la bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$. Queda completamente determinada por las n imágenes $L(\mathbf{e}_j)$ de los vectores de la base canónica. Se emplea también la notación

$$A = [L]_{B_c}^{B_c}$$

o simplemente $A = [L]_{B_c}$. La representación de L con respecto a otras bases será una matriz diferente, pero también de $m \times n$.

Ejemplos 5.4

1. Sea $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$, definida por $L(\mathbf{x}) = \begin{pmatrix} x_1 + x_2 \\ x_2 + x_3 \end{pmatrix}$ (ejemplo anterior). Tenemos

$$\mathbf{a}_1 \;=\; L(\mathbf{e}_1) = L\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{a}_2 = L\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathbf{a}_3 = L\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Por lo tanto,

$$A = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

Verificación:

$$A\mathbf{x} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_2 + x_3 \end{pmatrix} = L(\mathbf{x})$$

2. Dada A de $m \times n$, consideremos la transformación lineal $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ dada por

$$L(\mathbf{x}) = A\mathbf{x}$$

Es fácil ver que

$$L(\mathbf{e}_j) = A\mathbf{e}_j = \begin{pmatrix} a_{11} & \ldots & a_{1j} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mj} & \ldots & a_{mm} \end{pmatrix}\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} = \mathbf{a}_j$$

con $\mathbf{a}_j$ **la columna j-ésima de** A, por lo que la representación matricial de L en las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$ es precisamente A:

$$[L]_{B_c}^{B_c} = (L(\mathbf{e}_1), \ldots, L(\mathbf{e}_2)) = A$$

Por ejemplo, si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$,

$$L(\mathbf{x}) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{pmatrix}$$

y entonces

$$\begin{aligned} L(\mathbf{e}_1) &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix} = \mathbf{a}_1 \\ L(\mathbf{e}_2) &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} b \\ d \end{pmatrix} = \mathbf{a}_2 \end{aligned}$$

por lo que $(\mathbf{a}_1, \mathbf{a}_2) = A$.

El problema 5.2.3 implica entonces que la imagen $\mathrm{Im}(L)$ de una transformación lineal $L : \mathbb{R}^n \to \mathbb{R}^m$ es el espacio columna de $A = [L]_{B_c}^{B_c}$ (es decir, el espacio generado por los vectores columna de A) mientras que el núcleo $\mathrm{Nu}(L)$ es el espacio nulo de A.

3. Rotación general en $\mathbb{R}^2$. Tomemos la transformación $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, tal que a cada vector $\mathbf{x}$ lo hace rotar un ángulo θ, en sentido antihorario:

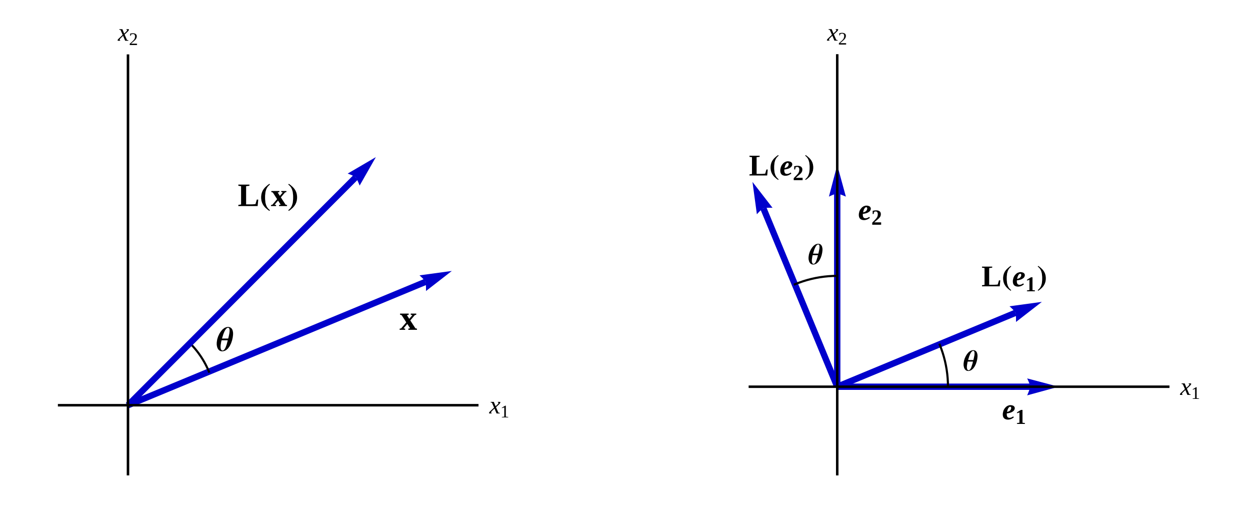


Figura 5.8: Rotación de ángulo θ.

Tenemos

$$L(\mathbf{e}_1) = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \quad \text{y} \quad L(\mathbf{e}_2) = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$$

Por lo tanto

$$A = (L(\mathbf{e}_1), L(\mathbf{e}_2)) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Entonces, para rotar un ángulo θ un vector $\mathbf{x}$ de $\mathbb{R}^2$ en sentido antihorario, se debe multiplicar A por $\mathbf{x}$:

$$\mathbf{y} = L(\mathbf{x}) = A\mathbf{x} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \cos\theta - x_2 \sin\theta \\ x_1 \sin\theta + x_2 \cos\theta \end{pmatrix}$$

Problemas 5.4

1. Hallar la matriz que representa, respecto de las bases canónicas, las siguientes transformaciones lineales y escribir la forma explícita de $L(\mathbf{x})$:
   (a) La aplicación $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ que representa una dilatación con factor c_1 a lo largo del eje x y c_2 a lo largo del eje y.
   (b) La reflexión $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ respecto de la recta $y = x$.
   (c) La rotación $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ de ángulo $\pi/4$ antihorario.
   (d) La rotación $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ de ángulo θ antihorario alrededor del eje z.
   (e) Idem anterior pero alrededor del eje y.
   (f) La proyección $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ sobre el plano xy.
   (g) $L : \mathbb{R}^2 \to \mathbb{R}^2$ lineal, definida por $L\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix}$, $L\begin{pmatrix}0\\2\end{pmatrix} = \begin{pmatrix}-2\\2\end{pmatrix}$.
   (h) $L : \mathbb{R}^3 \to \mathbb{R}^2$ lineal, definida por $L\begin{pmatrix}1\\0\\0\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix}$, $L\begin{pmatrix}1\\1\\0\end{pmatrix} = \begin{pmatrix}1\\-1\end{pmatrix}$, $L\begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$.
   (i) Determinar la imagen y núcleo de las transformaciones anteriores.

2. Mostrar, determinando las imágenes $L(\mathbf{e}_1)$ y $L(\mathbf{e}_2)$, que la siguiente matriz

$$A = \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix}$$

   representa en la base canónica la reflexión $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ respecto de la recta $y = mx$ que forma un ángulo θ con el eje x ($m = \tan\theta$). Verificar resultados previos.

5.4.1. Caso general

Consideremos una transformación lineal general $L : V \longrightarrow W$ entre espacios vectoriales V y W de dimensión n y m respectivamente. Sean

$$B_V = (\mathbf{v}_1, \ldots, \mathbf{v}_n), \qquad B_W = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$$

bases ordenadas de V y W. Cualquier vector $\mathbf{v} \in V$ puede escribirse en términos de los vectores de la base B_V:

$$\mathbf{v} = x_1\mathbf{v}_1 + \ldots + x_n\mathbf{v}_n$$

siendo $[\mathbf{v}]_{B_V} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \mathbf{x}$ el vector de coordenadas de $\mathbf{v}$ con respecto a la base B_V.

Por lo tanto, para L lineal,

$$L(\mathbf{v}) = x_1 L(\mathbf{v}_1) + \ldots + x_n L(\mathbf{v}_n)$$

Si ahora escribimos $L(\mathbf{v})$ y $L(\mathbf{v}_j)$ en términos de los vectores de la base B_W,

$$\begin{array}{rcl} L(\mathbf{v}) & = & y_1\mathbf{w}_1 + \ldots + y_m\mathbf{w}_m \\ L(\mathbf{v}_j) & = & a_{1j}\mathbf{w}_1 + \ldots + a_{mj}\mathbf{w}_m, \quad j = 1, \ldots, m \end{array}$$

tal que sus vectores de coordenadas en la base B_W son

$$[L(\mathbf{v})]_{B_W} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \mathbf{y}\,, \quad [L(\mathbf{v}_j)]_{B_W} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} = \mathbf{a}_j$$

obtenemos

$$\begin{array}{rcl} \mathbf{y} = [L(\mathbf{v})]_{B_W} & = & x_1[L(\mathbf{v}_1)]_{B_W} + \ldots + x_n[L(\mathbf{v}_n)]_{B_W} \\ & = & x_1\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \ldots + x_n\begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \\ & = & \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ & = & A\mathbf{x} \end{array}$$

es decir,

$$[L(\mathbf{v})]_{B_W} \quad = \quad A[\mathbf{v}]_{B_V}$$

donde A es la matriz de $m \times n$

$$A = ([L(\mathbf{v}_1)]_{B_W}, \ldots, [L(\mathbf{v}_n)]_{B_W}) = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}$$

Esta matriz A se denomina **representación matricial de L respecto de las bases B_V de V y B_W de W**. La denotaremos también como $A = [L]_{B_W}^{B_V}$. El resultado se puede resumir en el esquema gráfico de la Fig. (5.9).

Ejemplos 5.4.1

1. Sea $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ la transformación lineal definida por

   $$L(\mathbf{x}) = x_1\mathbf{b}_1 + (x_2 + x_3)\,\mathbf{b}_2$$

   donde $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$ y

   $$\mathbf{b}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \mathbf{b}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

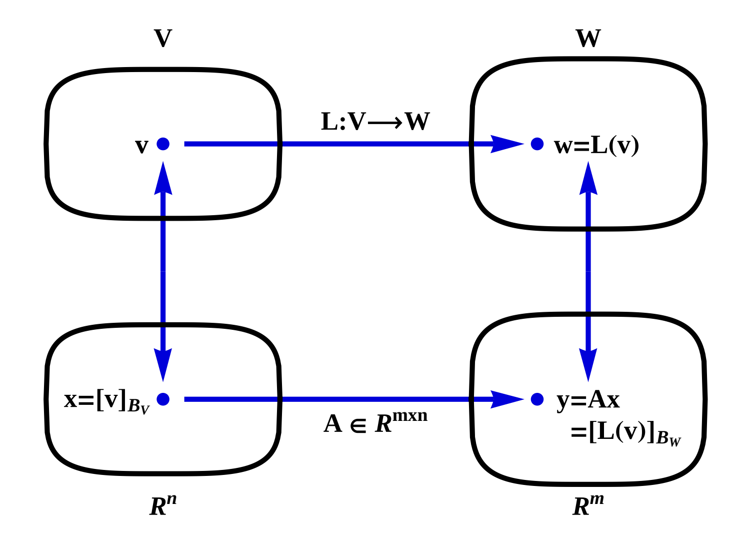


Figura 5.9: Esquema de una transformación lineal y su representación matricial.

Tenemos

$$\begin{aligned} L(\mathbf{e}_1) &= \mathbf{b}_1 = 1\mathbf{b}_1 + 0\mathbf{b}_2 \\ L(\mathbf{e}_2) = L(\mathbf{e}_3) &= \mathbf{b}_2 = 0\mathbf{b}_1 + 1\mathbf{b}_2 \end{aligned}$$

Por lo tanto, la representación matricial de L con respecto a las bases $B_c = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ de $V = \mathbb{R}^3$ y $B_W = (\mathbf{b}_1, \mathbf{b}_2)$ de $W = \mathbb{R}^2$ es

$$A = [L]_{B_W}^{B_c} = ([L(\mathbf{e}_1)]_{B_W}), [L(\mathbf{e}_2)]_{B_W}, [L(\mathbf{e}_3)]_{B_W}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

Así,

$$[L(\mathbf{x})]_{B_W} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 + x_3 \end{pmatrix}$$

que implica justamente $L(\mathbf{x}) = x_1\mathbf{b}_1 + (x_2 + x_3)\mathbf{b}_2$.

En cambio, en la base canónica de $\mathbb{R}^2$ obtenemos (¡probar!).

$$A_c = [L]_{B_c}^{B_c} = (L(\mathbf{e}_1), L(\mathbf{e}_2), L(\mathbf{e}_3)) = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

con

$$[L(\mathbf{x})]_{B_c} = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 - x_3 \\ x_1 + x_2 + x_3 \end{pmatrix}$$

es decir, $L(\mathbf{x}) = (x_1 - x_2 - x_3)\mathbf{e}_1 + (x_1 + x_2 + x_3)\mathbf{e}_2$, que coincide con $x_1\mathbf{b}_1 + (x_2 + x_3)\mathbf{b}_2$.

2. Sea $D : P_2 \to P_2$ el operador derivada $D = \frac{d}{dx}$ en el espacio de polinomios de grado ≤ 2. Como base ordenada de P_2 consideramos $B = (1, x, x^2)$. Tenemos

$$D(1) = 0,\ \ D(x) = 1,\ \ D(x^2) = 2x$$

Por lo tanto, su representación matricial en la base B es

$$A = [D]_B^B = ([D(1)]_B, [D(x)]_B, [D(x^2)]_B) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

Para $p(x) = a_0 + a_1 x + a_2 x^2 \in P_2$ tenemos $[p]_B = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$ y entonces

$$[D(p)]_B = A[p]_B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} a_1 \\ 2a_2 \\ 0 \end{pmatrix}$$

que implica $D(p) = a_1 + 2a_2 x + 0x^2 = a_1 + 2a_2 x$. Al ser P_2 de dimensión 3, todo operador lineal en P_2 puede entonces ser representado por una matriz de 3×3, tal como un operador en $\mathbb{R}^3$.

Notar que el espacio nulo de A es $\left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle$, que corresponde al subespacio de los polinomios de grado 0 ($p(x) = a_0$), es decir, al núcleo de D, mientras que su espacio columna es $\left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle$, que corresponde al subespacio de polinomios de grado 0 y 1, es decir, a la imagen de D. Esta correspondencia se discute a continuación.

Relación entre las propiedades de L y la matriz de representación

Sea $L : V \to W$ una transformación lineal, con $\dim V = n$, $\dim W = m$, y

$$A = [L]_{B_W}^{B_V}$$

la matriz que la representa con respecto a bases B_V de V y B_W de W. El rango y la nulidad de A **no dependen** de las bases elegidas, pues coinciden, respectivamente, con la dimensión de la imagen y del núcleo de L (los que no dependen de la representación).

En efecto, los n vectores columna $\mathbf{a}_j \in \mathbb{R}^m$ de A son las coordenadas de las n imágenes $L(\mathbf{v}_j) \in W$ en la base B_W. Como la correspondencia entre vectores y sus coordenadas en una base es un isomorfismo (Problema 5.3.5), la dimensión del subespacio de $\mathbb{R}^m$ generado por las n columnas de A (es decir, el rango de A) **es la misma** que la del subespacio de W generado por las n imágenes $L(\mathbf{v}_j)$, es decir, la misma que la dimensión de la imagen $\text{Im}(L)$. Análogamente, los vectores $\mathbf{x} \in \mathbb{R}^n$ del espacio nulo de A ($A\mathbf{x} = \mathbf{0}$) son las coordenadas de los vectores $\mathbf{v} \in V$ del núcleo $\text{Nu}(L)$ ($L(\mathbf{v}) = \mathbf{0}_W$), y por lo tanto la dimensión del espacio nulo de A (la nulidad de A) coincidirá con la del núcleo de L.

Se cumple entonces que

$$\text{rango}(A) = \dim \text{Im}(L)\,, \quad \text{nulidad}(A) = \dim \text{Nu}(L)$$

para **cualquier representación matricial** $A = [L]_{B_V}^{B_W}$ de L. Por lo tanto, la relación

$$\dim \text{Im}(L) + \dim \text{Nu}(L) = \dim V$$

resulta equivalente a

$$\text{rango}\,(A) + \text{nulidad}\,(A) = n$$

Los vectores $\in \mathbb{R}^m$ de una base del espacio columna de A son las coordenadas en la base B_W de los vectores de una base de la imagen $\text{Im}(L)$, mientras que los vectores $\in \mathbb{R}^n$ de una base del espacio nulo de A son las coordenadas en la base B_V de los vectores de una base del núcleo $\text{Nu}(L)$. Por lo tanto, pueden obtenerse bases de estos espacios mediante los métodos ya conocidos para obtener bases de los espacios columna y nulo de una matriz. Y la ecuación $L(\mathbf{v}) = \mathbf{w}$ es equivalente al sistema lineal $A[\mathbf{v}]_{B_V} = [\mathbf{w}]_{B_W}$.

Problemas 5.4 (continuación)

1. 3. A partir de consideraciones geométricas, encontrar la representación matricial $[L]_B^B$ respecto de la base $B = \left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}\right\}$ de la transformación lineal $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ que refleja todo vector respecto de la recta $y = x$. Mostrar que dicha matriz es diagonal. Comparar con la matriz que la representa en la base canónica $[L]_{B_c}^{B_c}$.

2. 4. Sea $D : P_3 \to P_3$ el operador derivada en el espacio de polinomios de grado ≤ 3.
   a) Determine su representación matricial A en la base $B = \{1, x, x^2, x^3\}$.
   b) Halle el núcleo y la imagen de D en este espacio, y su relación con el espacio columna y nulo de A.

5.5. Cambio de base

Determinaremos ahora en forma explícita la relación entre las representaciones matriciales de una transformación lineal L en distintas bases. La ventaja de "cambiar de base" es la posibilidad de encontrar una base en la que la matriz representativa de L sea simple (por ejemplo, diagonal) y por lo tanto sus propiedades esenciales (rango, imagen, etc.) resulten evidentes.

Consideremos primero un operador lineal $L : V \to V$, con V de dimensión finita n. Recordemos que si $B = \mathbf{v}_1, \ldots, \mathbf{v}_n$ y $B' = \mathbf{v}'_1, \ldots, \mathbf{v}'_n$ son dos bases de V, podemos escribir cualquier vector $\mathbf{v} \in V$ en las formas

$$\begin{aligned}\mathbf{v} &= x_1\mathbf{v}_1 + \ldots + x_n\mathbf{v}_n \\ &= x'_1\mathbf{v}'_1 + \ldots + x'_n\mathbf{v}'_n\end{aligned}$$

Las coordenadas $\mathbf{x} = [\mathbf{v}]_B = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ y $\mathbf{x}' = [\mathbf{v}]_{B'} = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$ en estas bases se relacionan por

$$\mathbf{x}' = S^{-1}\mathbf{x}$$

o en forma equivalente,

$$\mathbf{x} = S\mathbf{x}'$$

donde

$$S = \big([\mathbf{v}'_1]_B, \quad \ldots, \quad [\mathbf{v}'_n]_B\big)$$

es la matriz de cambio de base (o matriz de transición), de $n \times n$ y no singular, formada por las coordenadas de los vectores de la base B' en la base B. Por lo tanto, escribiendo

$$\begin{aligned} L(\mathbf{v}) &= y_1\mathbf{v}_1 + \ldots + y_n\mathbf{v}_n \\ &= y'_1\mathbf{v}'_1 + \ldots + y'_n\mathbf{v}'_n \end{aligned}$$

con $\mathbf{y} = [L(\mathbf{v})]_B = A\mathbf{x} = A[\mathbf{v}]_B$, $\mathbf{y}' = [L(\mathbf{v})]_{B'} = A'\mathbf{x}' = A'[\mathbf{v}]_{B'}$, obtenemos

$$\begin{aligned} \mathbf{y}' &= S^{-1}\mathbf{y} \\ &= S^{-1}A\mathbf{x} \\ &= S^{-1}AS\mathbf{x}' \end{aligned}$$

es decir,

Si $A = [L]_B^B$ es la representación de L en la base B y $A' = [L]_{B'}^{B'}$ su representación en la base B', las matrices A y A' se relacionan por

$$A' = S^{-1}AS$$

donde $S = \big([\mathbf{v}'{}_1]_B \quad \ldots \quad [\mathbf{v}'_n]_B\big)$ es la matriz de cambio de base.
Notar que si se conoce A', la relación anterior permite obtener A como (¡probar!)

$$A = SA'S^{-1}$$

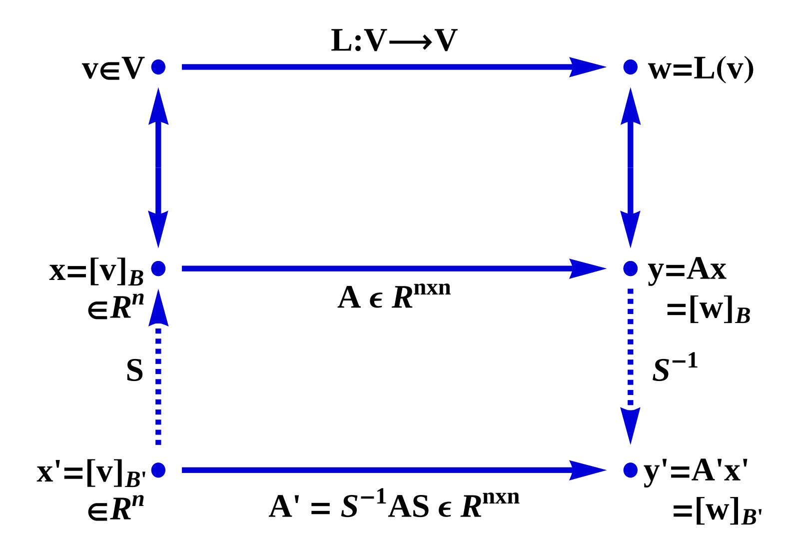


Figura 5.10: Esquema del cambio de base para la representación matricial.

Ejemplo 5.5.1 Consideremos nuevamente la reflexión $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ respecto de la recta $y = x$. En la base $B' = \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}-1\\1\end{pmatrix}\}$, formada por el vector $\mathbf{v}'_1 = \begin{pmatrix}1\\1\end{pmatrix}$ perteneciente a la recta y el vector $\mathbf{v}'_2 = \begin{pmatrix}-1\\1\end{pmatrix}$ perpendicular a la recta, la representacion es obvia, ya que $L(\mathbf{v}'_1) = \mathbf{v}'_1$ mientras que $L(\mathbf{v}'_2) = -\mathbf{v}'_2$ (véase problema 5.4.3):

$$A' = \begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}$$

Para hallar la representación matricial A de L en la base canónica $B = \{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\}$, determinamos primero la correspondiente matriz de cambio de base,

$$S = ([\mathbf{v}'_1]_B, [\mathbf{v}'_2]_B) = \begin{pmatrix}1 & -1\\1 & 1\end{pmatrix}$$

y su inversa $S^{-1} = \frac{1}{2}\begin{pmatrix}1 & 1\\-1 & 1\end{pmatrix}$. Luego

$$\begin{aligned} A = SA'S^{-1} &= \begin{pmatrix}1 & -1\\1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}\begin{pmatrix}1/2 & 1/2\\-1/2 & 1/2\end{pmatrix}\\ &= \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}\end{aligned}$$

Este resultado coincide con el hallado previamente (problema 5.4.1 (b)).

5.5.1. Matrices semejantes

Dos matrices A y A' de $n \times n$ son **semejantes** si y sólo si existe una matriz no singular S tal que

$$A' = S^{-1}AS$$

Si A' es semejante a A, entonces, multiplicando la igualdad anterior por S a izquierda y por S^{-1} a derecha, obtenemos

$$\begin{aligned} SA'S^{-1} &= S(S^{-1}AS)S^{-1} = (SS^{-1})A(SS^{-1}) \\ &= A \end{aligned}$$

es decir, $A = R^{-1}A'R$, con $R = S^{-1}$, por lo que A es también semejante a A'.

Por lo tanto, las matrices A y A' que representan a un operador lineal L en dos bases distintas son **semejantes**.

Dos propiedades fundamentales sobre matrices semejantes son las siguientes:

1. Las matrices semejantes tienen el mismo determinante:

$$\det(A') = \det(A)$$

   En efecto,

$$\begin{aligned} \det(A') &= \det(S^{-1}AS) \\ &= \det(S^{-1})\det(A)\det(S) \\ &= (\det(S))^{-1}\det(A)\det(S) \\ &= \det(A) \end{aligned}$$

Esto implica que el determinante $\det(A)$ **es una propiedad del operador lineal** L **representado por** A, permaneciendo **invariante** frente a cambios de base.

Así, comprobamos en el ejemplo 5.5.1 que $\det(A') = \det(A) = -1$.

2. Las matrices semejantes tienen la misma traza:

$$\mathrm{Tr}\,A' = \mathrm{Tr}\,A$$

donde la traza $\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{ii}$ es la suma de los elementos diagonales.

En efecto, probemos primero que para matrices A de $n \times m$ y B de $m \times n$, se cumple

$$\mathrm{Tr}\,(AB) = \mathrm{Tr}\,(BA)$$

Demostración:

$$\mathrm{Tr}AB = \sum_{i=1}^{n}(AB)_{ii} = \sum_{i=1}^{n}(\sum_{j=1}^{m} a_{ij}b_{ji}) = \sum_{j=1}^{m}(\sum_{i=1}^{n} b_{ji}a_{ij}) = \sum_{j=1}^{m}(BA)_{jj} = \mathrm{Tr}BA$$

Luego

$$\mathrm{Tr}\,A' = \mathrm{Tr}\,(SAS^{-1}) = \mathrm{Tr}\,((AS^{-1})S) = \mathrm{Tr}\,(A(S^{-1}S)) = \mathrm{Tr}\,A$$

Este resultado implica que la traza es también una propiedad del operador L representado por A, permaneciendo **invariante** frente a cambios de base.

Así, comprobamos en el ejemplo 5.5.1 que $\mathrm{Tr}\,(A') = \mathrm{Tr}(A) = 0$.

Finalmente, mencionemos una tercera propiedad relacionada con 1.:

3. Si A y A' son matrices semejantes de $n \times n$ e I_n es la matriz identidad, entonces

$$\det\,(A' - \lambda I_n) = \det(A - \lambda I_n)$$

para cualquier escalar λ.

Este determinante (que es un polinomio de grado n en λ denominado **polinomio característico**) juega un rol central en la teoría de autovalores, como veremos luego.
Demostración:

$$\begin{aligned}\det(A' - \lambda I_n) &= \det(S^{-1}A'S - \lambda I_n) = \det[S^{-1}(A - \lambda I_n)S] \\ &= (\det(S^{-1})\det(A - \lambda I_n)\det(S) \\ &= \det(A - \lambda I_n)\end{aligned}$$

Así, comprobamos en el ejemplo 5.5.1 que

$$\begin{aligned}\det(A' - \lambda I_2) &= \begin{vmatrix}1-\lambda & 0 \\ 0 & -1-\lambda\end{vmatrix} = (1-\lambda)(-1-\lambda) = \lambda^2 - 1 \\ \det(A - \lambda I_2) &= \begin{vmatrix}-\lambda & 1 \\ 1 & -\lambda\end{vmatrix} = \lambda^2 - 1 = \det(A' - \lambda I_2)\end{aligned}$$

Problemas 5.5

1. Dada $L:\mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definida por

$$L(\mathbf{x}) = A\mathbf{x}\,,\quad A = \begin{pmatrix}2 & 2 & 0\\ 1 & 1 & 2\\ 1 & 1 & 2\end{pmatrix}$$

   a) Obtener la expresión explícita de $L(\mathbf{x})$ para un vector $\mathbf{x} = \begin{pmatrix}x_1\\ x_2\\ x_2\end{pmatrix}$.

   b) Hallar su representación A' en la base $B' = \left\{\begin{pmatrix}1\\ -1\\ 0\end{pmatrix}, \begin{pmatrix}-2\\ 1\\ 1\end{pmatrix}, \begin{pmatrix}1\\ 1\\ 1\end{pmatrix}\right\}$.
   Verificar que $L(\mathbf{v}'_1) = 0$, $L(\mathbf{v}'_2) = \mathbf{v}'_2$, y $L(\mathbf{v}'_3) = 4\mathbf{v}'_3$, y que por lo tanto, A' es diagonal.
   c) Usar esta representación diagonal para hallar su núcleo e imagen.
   d) Usar esta representación diagonal para interpretar L geométricamente.
   e) Indique si el vector $\mathbf{v} = \begin{pmatrix}0\\ 1\\ 1\end{pmatrix}$ pertenece a la imagen de L.
   f) Verifique que la traza y determinante permanecen invariantes frente al cambio de base: $\mathrm{Tr}(A) = \mathrm{Tr}(A')$, $\det(A) = \det(A')$.

2. Si la matriz A del problema anterior representa ahora una transformación $L: P_2 \to P_2$ en la base canónica de P_2, con P_2 el espacio de polinomios reales de grado ≤ 2 (y su base canónica dada por $\{1, x, x^2\}$), dé una expresión para $L(p)$ y determine la base B' de P_2 donde la representación A' es diagonal (utilice resultados previos).

3. Considere la reflexión $L: \mathbb{R}^2 \to \mathbb{R}^2$ respecto de una recta que pasa por el origen y que forma un ángulo θ con el eje x, dada por $y = mx$, con $m = \tan\theta$ (graficar).
   a) Determine, a partir de consideraciones geométricas, su representación matricial A' en la base $B' = \left\{\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}, \begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix}\right\}$ formada por un vector perteneciente a la recta y otro perpendicular a dicha recta (¡verificar!). Muestre que A' es diagonal.
   b) Utilice a) y cambio de base para obtener la representación matricial A en la base canónica. Verifique que se obtiene el resultado del problema 5.4.2.

4. Sea $L: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida por $L\begin{pmatrix}1\\ 0\end{pmatrix} = \begin{pmatrix}2\\ 2\end{pmatrix}$, $L\begin{pmatrix}0\\ 1\end{pmatrix} = \begin{pmatrix}1\\ 1\end{pmatrix}$.
   a) Encuentre su representación matricial en la base canónica de $\mathbb{R}^2$ y dé una expresión para $L(\mathbf{x})$.
   b) Encuentre su representación matricial en la base $B' = \{\begin{pmatrix}1\\ -2\end{pmatrix}, \begin{pmatrix}1\\ 1\end{pmatrix}\}$, y verifique que es diagonal. c) Determine su núcleo e imagen.

5. a) Muestre que la representación matricial $A = [I]_B^B$ del operador identidad $I: V \to V$ con respecto a una base arbitraria B de un espacio vectorial V de dimensión n, es

siempre la matriz identidad I_n.
b) Muestre que la representación matricial del operador nulo $0 : V \to V$ en cualquier base B de V es la matriz nula de $n \times n$.

6. Muestre que si una matriz B es semejante a una matriz A, entonces:
i) B^2 es semejante a A^2.
ii) Si A no singular, B es no singular y B^{-1} es semejante a A^{-1}.

7. Dada $L : V \to W$ y A su representación matricial respecto de bases B_V de V y B_W de W (con dim$V = n$, dim$W = m$) muestre que su representación matricial A' respecto de bases B'_V de V y B'_W de W es

$$A' = R^{-1}AS$$

con $S = ([\mathbf{v}'_1]_{B_V}, \ldots, [\mathbf{v}'_n]_{B_V})$ la matriz de cambio de base en V (de $n \times n$, no singular) y $R = ([\mathbf{v}'_1]_{B_W}, \ldots, [\mathbf{v}'_n]_{B_W})$ la matriz de cambio de base en W (de $m \times m$, no singular).

5.6. Composición de transformaciones (operaciones sucesivas)

Consideremos dos transformaciones lineales $L : V \to W$, $G : W \to U$, con V, W, U espacios vectoriales. Supongamos que primero aplicamos L sobre un vector $\boldsymbol{v} \in V$, y luego aplicamos G sobre el vector resultante $\boldsymbol{w} = L(\boldsymbol{v}) \in W$. El resultado es el vector

$$\boldsymbol{u} = (G\,L)(\boldsymbol{v}) = G(L(\boldsymbol{v})) \tag{5.1}$$

que pertenece al espacio vectorial U:

$$\boldsymbol{v} \underset{L}{\to} \boldsymbol{w} \underset{G}{\to} \boldsymbol{u}$$

En (5.1), $G\,L : V \to U$ denota la *composición* de G con L (también escrita como $G \circ L$), que es también una transformación **lineal**.

Problemas 5.6.1

1. Demostrar que si $L : V \to W$ y $G : W \to U$ son transformaciones lineales, entonces $G\,L : V \to U$ definida por (5.1) es también una transformación lineal (probar que satisface $(G\,L)(\alpha\boldsymbol{v}_1 + \beta\boldsymbol{v}_2) = \alpha(G\,L)(\boldsymbol{v}_1) + \beta(G\,L)(\boldsymbol{v}_2)$).

2. Si $L : \mathbb{R}^3 \to \mathbb{R}^2$ está definida por $L(x, y, z) = (x + y, y + z)$,
y $G : \mathbb{R}^2 \to \mathbb{R}^2$ está definida por $G(x, y) = (x + y, 2x - 2y)$, encontrar una expresión para $(G\,L)(x, y, z) = G(L(x, y, z))$.

5.6.1. Representación matricial de la composición

Consideremos primero que $V=\mathbb{R}^n$, $W=\mathbb{R}^m$ y $U=\mathbb{R}^p$, y sean A_L (de $m\times n$) y A_G (de $p\times m$) las representaciones matriciales de L y G en las bases canónicas correspondientes, tal que $L(\boldsymbol{v})=A_L\boldsymbol{v}$, $G(\boldsymbol{w})=A_G\boldsymbol{w}$. Entonces

$$GL(\boldsymbol{v})=G(L(\boldsymbol{v}))=G(A_L\boldsymbol{v})=A_G(A_L\boldsymbol{v})=(A_GA_L)\boldsymbol{v}$$

por lo que la representación matricial de la composición $G\,L$ con respecto a las bases canónicas de V y U **es el producto** A_GA_L de las matrices que representan a G y a L:

$$A_{GL}=A_G\,A_L \tag{5.2}$$

Observar que la matriz A_G se aplica a la izquierda de la matriz A_L. El producto está así bien definido.

Ejemplo 5.6: En el problema 5.6.2, en las bases canónicas de $\mathbb{R}^3$ y $\mathbb{R}^2$, obtenemos las representaciones matriciales (¡verificar!).

$$A_L=\begin{pmatrix}1&1&0\\0&1&1\end{pmatrix},\quad A_G=\begin{pmatrix}1&1\\2&-2\end{pmatrix}$$

tal que

$$\begin{aligned}L\begin{pmatrix}x\\y\\z\end{pmatrix}&=\begin{pmatrix}1&1&0\\0&1&1\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}x+y\\y+z\end{pmatrix}\\G\begin{pmatrix}x\\y\end{pmatrix}&=\begin{pmatrix}1&1\\2&-2\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}x+y\\2x-2y\end{pmatrix}\end{aligned}$$

Por lo tanto,

$$\begin{aligned}(GL)\begin{pmatrix}x\\y\\z\end{pmatrix}&=\begin{pmatrix}1&1\\2&-2\end{pmatrix}\begin{pmatrix}1&1&0\\0&1&1\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}\\&=\begin{pmatrix}1&2&1\\2&0&-2\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}\\&=\begin{pmatrix}x+2y+z\\2x-2z\end{pmatrix}\end{aligned} \tag{5.3}$$

o sea, $A_{GL}=A_GA_L=\begin{pmatrix}1&2&1\\2&0&-2\end{pmatrix}$, con $(G\,L)(x,y,z)=(x+2y+z,2x-2z)$.

En el caso general, para representaciones matriciales en bases arbitrarias B_V, B_W y B_U de V, W, U, se obtiene también (se deja la demostración para el lector)

$$[GL]_{B_U}^{B_V}=A_GA_L\ \text{ con }\ A_L=[L]_{B_W}^{B_V},\ \ A_G=[G]_{B_U}^{B_W}$$

Este resultado es válido para espacios vectoriales generales de dimensión finita.

5.6.2. Potencias de operadores lineales

Recordemos que si $W = V$, la transformación lineal $L : V \to V$ se denomina también operador lineal o endomorfismo. Queda en este caso definido el cuadrado $L^2 = L\,L$ como la composición de L con L:

$$L^2(\boldsymbol{v}) = L(L(\boldsymbol{v})) \tag{5.4}$$

Si A_L es la representación matricial de L en una cierta base de V, los resultados anteriores implican que la representación matricial de L^2 es el cuadrado de la matriz A_L:

$$A_{L^2} = A_L\,A_L = A_L^2 \tag{5.5}$$

En general, la potencia L^k $\forall\ k \geq 2$ queda definida por

$$L^k(\boldsymbol{v}) = L(L^{k-1}(\boldsymbol{v})),\quad k \geq 2$$

y su representación matricial es

$$A_{L^k} = A_L\,A_{L^{k-1}} = A_L^k$$

es decir, la potencia k de la matriz A_L. Por definición, $L^0 = I$, con I el operador identidad. Y si L es un operador lineal inversible (o sea un *automorfismo*: un isomorfismo de V en V), de forma que existe la transformación inversa L^{-1} (tal que $L^{-1}L = I_V$) entonces

$$A_{L^{-1}} = A_L^{-1} \tag{5.6}$$

ya que $A_{L^{-1}}A_L = I_n$. La representación matricial de la inversa L^{-1} es la inversa de la matriz A_L que representa a L, como se vió previamente.

Ejemplo: Si $L : \mathbb{R}^2 \to \mathbb{R}^2$ está dado por $L\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x+y \\ -x-2y \end{pmatrix}$, su representación matricial en la base canónica es (¡probar!)

$$A_L = \begin{pmatrix} 2 & 1 \\ -1 & -2 \end{pmatrix}$$

La representación matricial de L^2 es entonces

$$A_{L^2} = A_L^2 = \begin{pmatrix} 2 & 1 \\ -1 & -2 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ -1 & -2 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$$

de forma que

$$L^2\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3x \\ 3y \end{pmatrix} = 3\begin{pmatrix} x \\ y \end{pmatrix}$$

es decir, $L^2(\boldsymbol{x}) = 3\boldsymbol{x}$, lo que equivale a $L^2 = 3I$, con I el operador identidad.

Además, como A_L es no singular, L es inversible y la representación matricial de su inversa está dada por

$$A_{L^{-1}} = A_L^{-1} = \frac{1}{3}\begin{pmatrix} 2 & 1 \\ -1 & -2 \end{pmatrix}$$

de forma que

$$L^{-1}\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 2 & 1 \\ -1 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 2x+y \\ -x-2y \end{pmatrix}$$

o sea, $L^{-1} = L/3$. Se verifica entonces $L^{-1}L = I$.

Problemas 5.6.2

1. Si $L:\mathbb{R}^3\to\mathbb{R}^3$ queda definido por $L(x,y,z)=(y,x,-z)$,
   i) mostrar que en la base canónica,
$$A_L=\begin{pmatrix}0&1&0\\1&0&0\\0&0&-1\end{pmatrix}$$
   ii) Hallar A_L^2 y verificar que L^2 es el operador identidad.

2. Sea P_2 el subespacio de los polinomios reales de grado ≤ 2.
   Si $D:P_2\to P_2$ denota el operador derivada ($D=\frac{d}{dx}$),
   i) Hallar las representaciones matriciales A_D y A_{D^2} de D y la derivada segunda D^2 en las base canónica $\{1,x,x^2\}$ y verificar que
$$A_{D^2}=A_DA_D=A_D^2$$

5.6.3. Composición de transformaciones lineales en $\mathbb{R}^2$

Consideremos ahora algunos ejemplos de operadores lineales $L:\mathbb{R}^2\to\mathbb{R}^2$. La reflexión de un vector $\boldsymbol{v}$ respecto de la recta $y=x$ está dada por (¡recordar!)

$$L\binom{x}{y}=\begin{pmatrix}0&1\\1&0\end{pmatrix}\binom{x}{y}=\binom{y}{x}$$

siendo su representación matricial en la base canónica $A_L=\begin{pmatrix}0&1\\1&0\end{pmatrix}$.
Es evidente de la definición que $L^2\binom{x}{y}=L(L\binom{x}{y})=\binom{x}{y}$ $\forall$ x, y, o sea, $L^2=I$ (operador identidad), verificándose que (¡probar!)

$$A_L^2=\begin{pmatrix}1&0\\0&1\end{pmatrix}$$

Es, decir, la inversa de una reflexión L es la misma reflexión L, como es obvio geométricamente.

La rotación de ángulo $\pi/2$ en sentido antihorario de un vector $\boldsymbol{v}$ está dada por

$$R\binom{x}{y}=\begin{pmatrix}0&-1\\1&0\end{pmatrix}\binom{x}{y}=\binom{-y}{x}$$

siendo su representación matricial en la base canónica $A_R=\begin{pmatrix}0&-1\\1&0\end{pmatrix}$.

Consideremos ahora la transformación lineal $R\,L$ que primero refleja un vector respecto de

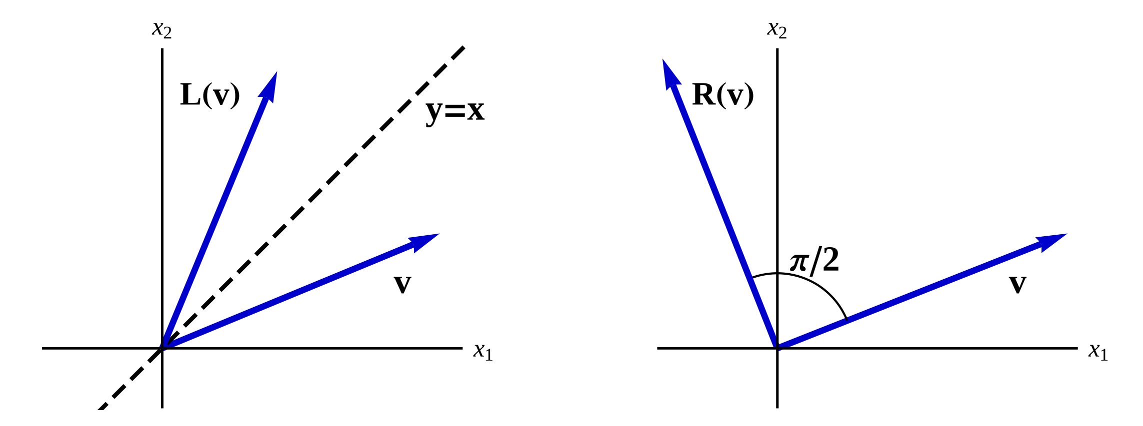


Figura 5.11: Reflexión respecto de la recta $y = x$ (izquierda) y rotación de ángulo $\pi/2$ antihorario (derecha).

la recta $y = x$ y luego lo rota un ángulo de $\pi/2$ en sentido antihorario. Su representación matricial en la base canónica será

$$A_{R\,L} = A_R\,A_L = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

y por lo tanto,

$$R\,L \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ y \end{pmatrix}$$

Esto representa una **reflexión respecto del eje** y (¡mostrar!).

Por otro lado, la transformación lineal $L\,R$, que primero rota un vector un ángulo de $\pi/2$ en sentido antihorario y luego lo refleja respecto de la recta $y = x$, queda representada por la matriz

$$A_{L\,R} = A_L\,A_R = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

y por lo tanto,

$$L\,R \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}$$

Esto representa una **reflexión respecto del eje** x (¡mostrar!).

Por lo tanto, vemos que el resultado final depende **del orden de la composición**, lo que se refleja en la **no conmutatividad** del producto matricial asociado.

Se define el **conmutador** de dos operadores $L : V \to V$, $G : V \to V$ como

$$[G, L] = GL - LG$$

que es en general un operador no nulo. La matriz que lo representa en una base de V es el conmutador de las matrices que representan a G y L en dicha base:

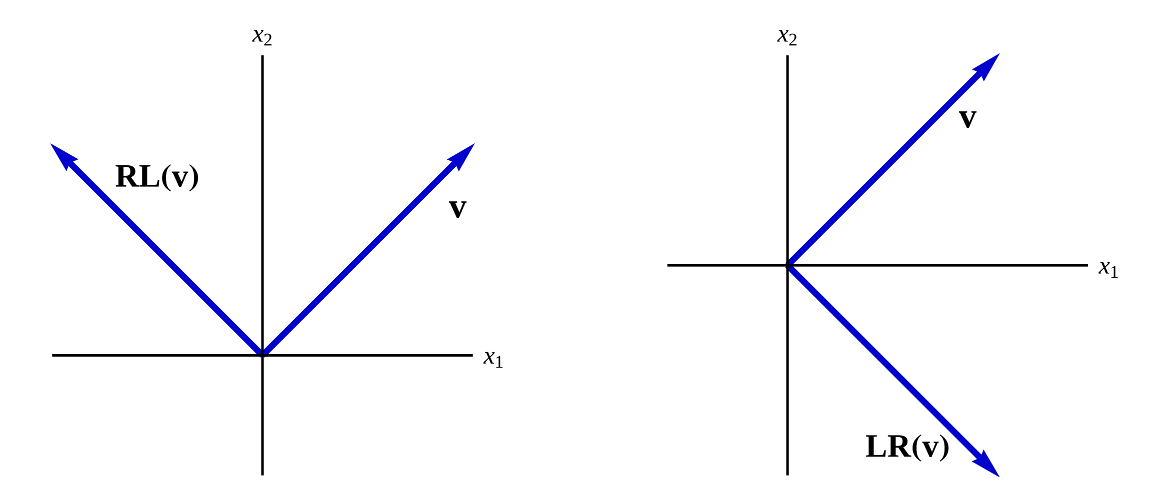


Figura 5.12: No conmutatividad de la composición de la reflexión y rotación.

$A_{[G,L]} = A_G\, A_L - A_L\, A_G$. En el ejemplo anterior se obtiene

$$A_R\, A_L - A_L\, A_R = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = 2\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

que implica $RL - LR = 2RL$, es decir, $LR = -RL$.

Recordemos finalmente que la proyección ortogonal de un vector sobre el eje x está dada por (¡graficar!)

$$P_x \binom{x}{y} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \binom{x}{y} = \binom{x}{0}$$

y la proyección ortogonal de un vector sobre el eje y está dada por

$$P_y \binom{x}{y} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \binom{x}{y} = \binom{0}{y}$$

Problemas 5.6.3

1. a) Hallar la representación matricial en la base canónica de la inversa de los operadores R y L anteriores.
   b) ¿Tienen P_x y P_y inversa?
   c) Encuentre la representación matricial de P_xP_y y P_yP_x. Justificar el resultado.

2. Sea $F : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ el operador lineal que primero rota todo vector un ángulo $\pi/2$ en sentido antihorario y luego lo proyecta sobre el eje x. Hallar su representación matricial en la base canónica y una expresión para $F\binom{x}{y}$.

3. Sea $G : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ el operador lineal que primero proyecta todo vector sobre el eje x y luego lo rota un ángulo $\pi/2$ en sentido antihorario. Hallar su representación matricial en la base canónica y una expresión para $G\binom{x}{y}$.

4. Recordando que la representación matricial en la base canónica de una reflexión respecto de la recta $y = mx$, con $m = \tan\theta$, es

$$\begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}$$

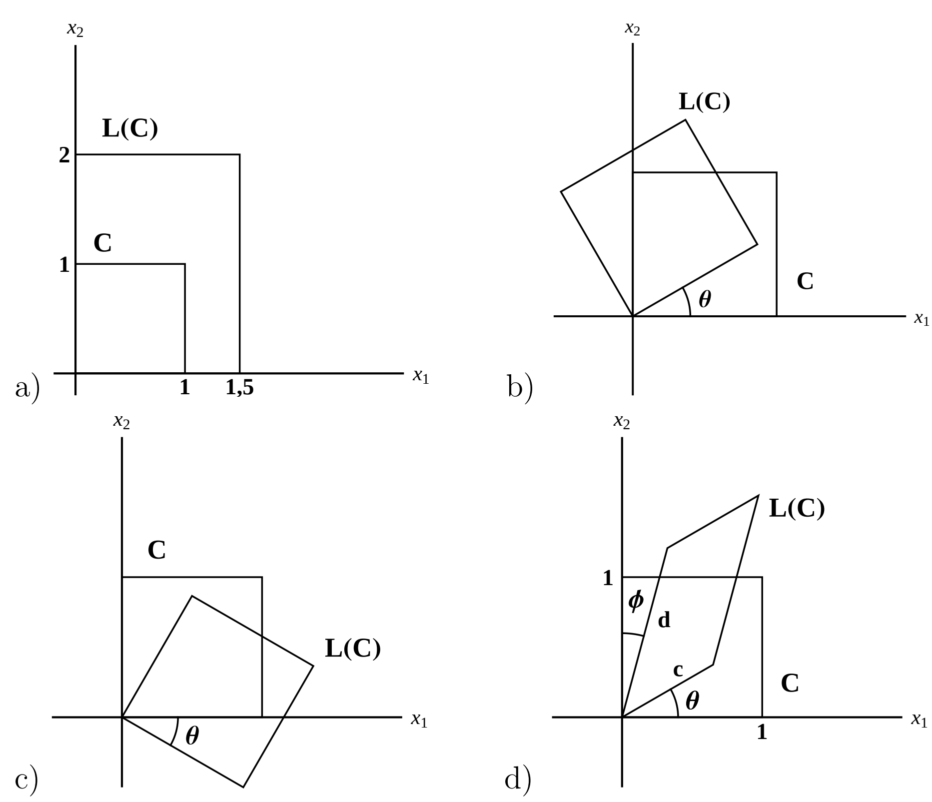
a) Halle su determinante.
b) Muestre que la composición de dos reflexiones es una rotación.

5. Recordando que la representación matricial en la base canónica de una rotación de ángulo θ en sentido antihorario es

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

a) Halle su determinante.
b) Muestre que la composición de dos rotaciones es otra rotación.
c) Muestre que la composición de una reflexión con una rotación es otra reflexión.

6. Encuentre, a partir de las figuras, la representación matricial en la base canónica de las siguientes transformaciones lineales $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$:



Bibliografía

[1] Grossman, S. (2008). *Algebra lineal,* 6a. ed. McGraw Hill.

[2] Strang, G. (2009). *Linear algebra and its applications*, 4th. ed., Wellesley Cambride press.

[3] Lay, D.C., Murrieta, J.M. (2007). *Algebra lineal y sus aplicaciones*, 3a. ed., Pearson.

[4] Kolman, B., Hill, D. (2006). *Algebra lineal*, 8a. ed., Pearson.

[5] Bru, J., Clement, J., Mas, J., Urbano, A. (2004). *Algebra lineal*, Alfaomega.

[6] Poole, D. (2007). *Algebra lineal, una introducción moderna*, 3a. ed., Thomson.

[7] Anton, H. (2001). *Introducción al Algebra Lineal,* Limusa.

[8] Hoffman, K., Kunze, R. (1971). *Linear Algebra*, Ed. Prentice-Hall Inc., New Jersey.

[9] Shilov, G.E. (1977). *Linear Algebra*, Ed. Dover Pub. New York.

Los autores

Costa, Viviana

Recibió el título de Lic. en Matemática en 1989 por la Universidad Nacional de La Plata (UNLP), Argentina, de MSc. en Simulación Numérica y Control en 2002 por la Universidad Nacional de Buenos Aires, Argentina y el título de Dra. en Enseñanza de las Ciencias en 2013 por la Universidad Nacional del Centro de la Provincia de Buenos Aires, Argentina. Desde el año 1986 se ha desempeñado como docente en diversos cursos de Matemática en Facultades de la UNLP. Sus intereses investigativos incluyen: control de sistemas dinámicos y educación en ciencias. Actualmente es Profesor con dedicación exclusiva en la Facultad de Ingeniería de la UNLP en la cátedra de Matemática C y Coordinador de la Unidad de Investigación, Desarrollo, Extensión y Transferencia: "Investigación en Metodologías Alternativas para la enseñanza de las Ciencias" (IMApEC).

Rossignoli, Raúl

Recibió el título de Lic. en Física en 1983 por la UNLP y el de Dr. en Física en 1987 en la Facultad de Ciencias Exactas de la UNLP. Desde 1983 se ha desempeñado como docente en diversas asignaturas de Matemática y Física en la UNLP, siendo actualmente Profesor Titular en la Facultad de Ingeniería (cátedra de Matemática C) y en la Facultad de Ciencias Exactas. También ha desarrollado actividades de investigación desde 1984 en Física Cuántica y Teoría de la Infor-

mación, siendo actualmente Investigador Principal de la Comisión de Investigaciones Científicas (CIC) y coordinador del grupo de Información cuántica y Sistemas cuánticos de muchos cuerpos en el Instituto de Física de La Plata dependiente del CONICET. Ha publicado más de 100 artículos en revistas científicas internacionales y es coautor de un libro sobre Ecuaciones Diferenciales. Dirige un proyecto de investigación en la Facultad de Ciencias Exactas e integra un proyecto de investigación sobre Enseñanza de las Ciencias Básicas en la Facultad de Ingeniería.

Sorichetti, Carlos

Recibió el título de Ingeniero en Telecomunicaciones en 1983 en la Facultad de Ingeniería de la UNLP y de Licenciado en Matemática en 2012 en la Facultad de Ciencias Exactas de la UNLP. Es docente graduado en la Facultad de Ingeniería desde 1983, desempeñando actualmente los cargos de Profesor Adjunto con Dedicación Exclusiva en las Cátedras de Matemática B y Matemática C y de Profesor Adjunto con Dedicación Simple en la Cátedra de Matemática D. También se ha desempeñado en Cátedras de Matemática de la Facultad de Ciencias Exactas de la UNLP. Recibió en 2013, la distinción *Dr. Joaquín V. González,* otorgada a los diez mejores promedios de los egresados de la Facultad de Ciencias Exactas de la UNLP del año 2012.

Vampa, Victoria

Recibió el título de Lic. en Matemática Aplicada en 1984 por la UNLP, de MSc. en Simulación Numérica y Control en 2004 por la Universidad Nacional de Buenos Aires y el título de Dra. de la Facultad de Ciencias Exactas, Área Matemática, en 2011 por la UNLP. Desde 1984 se ha desempeñado como docente en asignaturas de Matemática en Facultades de la UNLP. Ha desarrollado tareas de investigación en el área de Análisis Numérico y Matemática Aplica-

da, en particular en el desarrollo de herramientas que utilizan la Transformada Wavelet en la resolución numérica de ecuaciones diferenciales y en el procesamiento de señales. Actualmente es Profesor con dedicación exclusiva en la Facultad de Ingeniería de la UNLP en la cátedra de Matemática C.

Libros de Cátedra

Álgebra lineal con aplicaciones : parte I / Viviana Costa ... [et al.] ; coordinación general de Raúl Rossignoli. - 1a edición para el alumno - La Plata : Universidad Nacional de La Plata ; La Plata : EDULP, 2018.
Libro digital, PDF

Archivo Digital: descarga y online
ISBN 978-950-34-1603-7

1. Álgebra Lineal. I. Costa, Viviana II. Rossignoli, Raúl, coord.
CDD 512.5

Diseño de tapa: Dirección de Comunicación Visual de la UNLP

Universidad Nacional de La Plata – Editorial de la Universidad de La Plata
47 N.º 380 / La Plata B1900AJP / Buenos Aires, Argentina
+54 221 427 3992 / 427 4898
edulp.editorial@gmail.com
www.editorial.unlp.edu.ar

Edulp integra la Red de Editoriales Universitarias Nacionales (REUN)

Primera edición, 2018
ISBN 978-950-34-1603-7





